

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧЕРКАСЬКИЙ ДЕРЖАВНИЙ ТЕХНОЛОГІЧНИЙ УНІВЕРСИТЕТ

ВІСНИК

ЧЕРКАСЬКОГО ДЕРЖАВНОГО ТЕХНОЛОГІЧНОГО УНІВЕРСИТЕТУ

Технічні науки
Заснований у грудні 1996 р.

2/2022

Редакційна колегія:

Al-Azzeh J., д-р філософії, доцент (Йорданія)
Антонюк В. С., д-р техн. наук, професор
Артемчук В.О., д-р техн. наук, с.н.с.
Базіло К. В., д-р техн. наук, професор
Батраченко О. В., д-р техн. наук, доцент
Бондаренко М. О., д-р техн. наук, професор
Бондаренко Ю. Ю., канд. техн. наук, професор
Вамболь В. В., д-р техн. наук, професор
Ващенко В. А., д-р техн. наук, професор
Вітенько Т. М., д-р техн. наук, професор
Вербицький С. Б., канд. техн. наук
Вязовик В. М., д-р техн. наук, доцент
Гальченко В. Я., д-р техн. наук, професор
Гевод В. С., д-р хім. наук, доцент
Зайчук О. В., д-р техн. наук, професор
Заболотній С. В., д-р техн. наук, професор
Iavich M., д-р філософії, професор (Грузія)
Kažys R. J., д-р техн. наук, професор (Литва)
Коваленко І. Л., д-р техн. наук, професор
Лобода О. А., д-р хім. наук
Лукашенко В. М., д-р техн. наук, професор
Мусієнко М. П., д-р техн. наук, професор
Осипенко В. І., д-р техн. наук, професор
Палагін В. В., д-р техн. наук, професор
Півоваров О. А., д-р техн. наук, професор
Плахотний О. П., д-р техн. наук, доцент
Прокопенко Т. О., д-р техн. наук, професор
Ситник О. О., д-р техн. наук, професор
Скиба М. І., канд. техн. наук, доцент
Смотраєв Р. В., канд. техн. наук, доцент
Столярєнко Г. С., д-р техн. наук, професор
Сухенко В. Ю., д-р техн. наук, професор
Уткіна Т. Ю., канд. техн. наук, доцент
Фауре Е. В., д-р техн. наук, професор (гол. ред.)
Федоров Є. Є., д-р техн. наук, професор
Фролова Л. А., д-р техн. наук, професор
Хоменко О. М., канд. хім. наук, доцент
Яценко І. В., д-р техн. наук, професор

ЗАСНОВНИК І ВИДАВЕЦЬ – Черкаський
державний технологічний університет

Свідоцтво про державну реєстрацію
друкованого засобу масової інформації
КВ № 6061 від 16.04.2002.

Виходить 4 рази на рік

У ВИПУСКУ:

- ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ
- ХІМІЧНІ ТЕХНОЛОГІЇ
ТА ІНЖЕНЕРІЯ,
ЕКОЛОГІЧНА БЕЗПЕКА

Редакційна рада видання:

Григор О. О., д-р політ. наук, професор,
ректор Черкаського державного
технологічного університету (голова ради);
Гончаров А. В., канд. техн. наук, доцент,
перший проректор ЧДТУ;
Фауре Е. В., д-р техн. наук, професор,
проректор з науково-дослідної роботи
і міжнародних зв'язків ЧДТУ;
Столярєнко Г. С., д-р техн. наук, професор;
Литвин О. В., нач. ред.-видав. відділу ЧДТУ

АДРЕСА РЕДАКЦІЇ:

ЧДТУ, к. 604/1, бульвар Шевченка, 460,
м. Черкаси, 18006,
тел. (0472) 51-36-71

vtn.chdtu.edu.ua (подання статей)

Сисоєнко С. В. (відповідальний секретар)
e-mail: vtn@chdtu.edu.ua

2•2022

ВІСНИК

**Черкаського державного
технологічного університету**

ТЕХНІЧНІ НАУКИ

VISNYK

**Cherkaskogo derzhavnogo
tehnologichnogo universitetu**

TECHNICAL SCIENCE

2 • 2022

**Видання включене до Переліку
наукових фахових видань України
як фахове видання з технічних наук
(спеціальності 101, 113, 121, 122, 123, 125,
126, 131, 132, 133, 151, 152, 161, 172)
категорія «Б»
(накази МОНУ від 02.07.2020 №886,
від 24.09.2020 №1188)**

**Видання зареєстровано
у міжнародних наукометричних базах,
каталогах і системах пошуку:**
Index Copernicus Journals Master List,
BASE, CiteFactor, Google Scholar, Crossref,
Directory of Open Access Journals (DOAJ),
Eurasian Scientific Journal Index (ESJI),
ResearchBib – Academic Resource Index,
Ulrich's Periodicals Directory, WorldCat,
Open Ukrainian Citation Index

Електронна версія:
vtn.chdtu.edu.ua; visnyk.chdtu.edu.ua

Затверджено Вченою радою
Черкаського державного
технологічного університету,
протокол № 16 від 20.06.2022.

За повного або часткового
передрукування матеріалів
посилання на «Вісник Черкаського
державного технологічного
університету» є обов'язковим.

ЗМІСТ

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

<i>Prokopenko T. O., Pidkuiko O. I.</i> Devising a graph-analytical model for situation project management of scrum-based projects in the field of information technology.....	4
<i>Федоров Є. Є., Уткіна Т. Ю.</i> Метод кластеризації квазіперіодичного сигналу на основі алгоритму клонального відбору	11
<i>Скітер І. С., Савельєв М. В., Купріяничук С. В., Хоменко Д. О.</i> Алгоритм моніторингу потужностей критеріїв тестування експоненціальності щільності потоку нейтронів в «Об'єкті Укриття»	22
<i>Зінченко І. Г., Лавданська О. В.</i> Сучасні технології оцінювання ефективності діджиталізації.....	34
<i>Yakymenko D. O., Kataieva Ye. Yu.</i> Methods and means of intelligent analysis of text documents.....	43

ХІМІЧНІ ТЕХНОЛОГІЇ ТА ІНЖЕНЕРІЯ,
ЕКОЛОГІЧНА БЕЗПЕКА

<i>Кириченко Є. П.</i> Методика визначення критичних значень параметрів зовнішніх термічних дій на піротехнічні металооксидні вироби в умовах експлуатації	53
<i>Сушацький Ю. В., Шеніда М. В., Созанський М. А., Знак З. О., Хом'як С. В.</i> Кінетика окиснювальної деградації метилового оранжевого з використанням процесу “Сонопероксат”	64

CONTENTS

INFORMATION TECHNOLOGIES

<i>Prokopenko T. O., Pidkuiko O. I.</i> Devising a graph-analytical model for situation project management of scrum-based projects in the field of information technology	4
<i>Fedorov E. E., Utkina T. Yu.</i> Method of clusterization of quasiperiodic signal based on clonal selection algorithm.....	11
<i>Skiter I. S., Saveliev M. V., Kupriyanchuk S. V., Khomenko D. O.</i> Algorithm for monitoring the power of criteria for testing the exponentivity of neutron flux density in the Shelter Object	22
<i>Zinchenko I. H., Lavdanska O. V.</i> Modern technologies for evaluating the effectiveness of digitalization.....	34
<i>Yakymenko D. O., Kataieva Ye. Yu.</i> Methods and means of intelligent analysis of text documents	43

CHEMICAL TECHNOLOGIES AND ENGINEERING,
ECOLOGICAL SECURITY

<i>Kirichenko E. P.</i> Technique for determination of critical values of parameters of external thermal actions on pyrotechnical metal oxide products in operating conditions.....	53
<i>Sukhatskiy Yu. V., Shepida M. V., Sozanskyi M. A., Znak Z. O., Khomyak S. V.</i> Kinetics of oxidative degradation of methyl orange using the “Sonoperoxate” process.....	64

[0000-0002-6204-0708] **T. O. Prokopenko**, Dr. Tech. Sc., Professor,
e-mail: t.prokopenko@chdtu.edu.ua
[0000-0001-5722-4863] **O. I. Pidkuiko**

e-mail: o.i.pidkuiko.fitis20@chdtu.edu.ua

Черкаський державний технологічний університет
б-р Шевченка, 460, м. Черкаси, 18006, Україна

DEVISING A GRAPH-ANALYTICAL MODEL FOR SITUATION PROJECT MANAGEMENT OF SCRUM-BASED PROJECTS IN THE FIELD OF INFORMATION TECHNOLOGY

A graph-analytical model for situational project management in the field of information technology, which can ensure effective management decision-making, has been proposed. All features inherent in the implementation of the Scrum methodology used in the management of information technology projects have been taken into consideration. Ambiguous situations that constantly arise during the implementation of projects affect their effectiveness. Therefore, constant monitoring of the project efficiency will enable the project manager to make management decisions, which can contribute to improving the efficiency of the project. The impact of changing situational factors on the situation, which provides real-time monitoring of the situation for the project in Scrum conditions, has been studied. The model of situational management is built on the basis of the complex application of graph theory, situational analysis and expert methods, which ensure obtaining more accurate and reliable information for decision-making in conditions of rapidly changing circumstances.

Ключові слова: Scrum, Sprin, factors, situation, graph.

Introduction The trend of globalization and digitization of the modern world activates the development of innovative technologies, particularly in the management system. The best experience of applying such innovations, especially by companies in the field of information technology (IT), provides increased efficiency due to dual management in current and strategic activities. Therefore, it is important to implement the set tasks by applying the project approach, which is based on the Agile software development. In particular, Scrum is a framework that helps to increase the productivity of developers. However, ambiguous situations arise quite often. They are unpredictable in nature and affect the effectiveness of the project. As stated in [1-3], the basis of Agile software development is a short cycle that provides a minimum increase in the software product, i.e. iteration. Paper [4] indicates that the software development process requires constant adjustment of the project cost with direct control of quality, timing and duration. The use of Scrum provides opportunities for an innovative approach to the management system. However, within the framework of the iteration, the study of various situations that may arise and affect the course of the project is very important. According to Scrum [1-3], the dynamics of the imple-

mentation of an IT project requires a clear statement of tasks, a clear formulation of requirements, the distribution of processes between performers, a clear coordination of actions and decisions, and adaptation to the style of teamwork.

In the process of IT project management, it is necessary to ensure the flexibility of the software product development processes, adaptation to internal and external factors, based on the fact that both the environment and the conditions for the implementation of the IT project will change. However, in most cases it is difficult to predict how these changes will occur. Therefore, most IT companies during the implementation of projects choose to use the flexible Scrum methodology, which ensures both product quality control and the quality of the processes themselves.

In paper [5], the authors consider an IT project as a complex organizational and technological system, which is characterized by integrated management approaches, as well as strict standards and principles. The peculiarities of using the Scrum methodology in the IT project management system require taking into account the possibilities of various situations that can lead to negative consequences and affect the effectiveness of the IT project.

Situational management of complex organizational and technological objects focuses on the fact that the need for different management methods is determined by a specific situation. Situational analysis represents complex technologies of preparation, adoption and implementation of management decisions, which are based on the analysis of a separate management situation. In the course of situational management, methods of fuzzy logic [6], logic-linguistic models [7], training and generalization procedures are used in the generation of management decisions according to the current situation [8].

There are a number of methods that implement different ways of the situational approach. Paper [9] considers methods focused on the formation of an artificial language for describing situations and relations between objects. However, the construction of a semantic model of the IT project and the processes occurring in it is not always justified and is cumbersome when describing situations in the Scrum management system. The process of creating a language is quite time-consuming and does not always lead to a positive result.

Paper [10] reports the results of situation research based on a group of methods based on the representation of the situation in the form of a set of values of a fixed set of features. When solving certain problems, these methods can be combined, but it is necessary to take into account a number of significant differences.

Our review of works [4-10] has shown that the existing models and methods do not fully describe situations in Scrum and provide complete and accurate information for making management decisions in real time.

The use of situational analysis in IT project management under Scrum conditions has a number of advantages that contribute to obtaining more accurate, reliable and complete information for decision-making in a changing environment. Therefore, the development of a model for situational management of an IT project in the context of the Scrum methodology is an urgent scientific task. At the same time, for complex organizational and technical systems, it is necessary not only to identify the current situation and the set of management decisions corresponding to it, but also to determine rational ways of achieving the goals of the functioning of the IT project management system, for which it is necessary to calculate the possible consequences of the sequence of management decisions for several steps for-

ward. These tasks require the involvement of additional methods. The development of a graph-analytical model of IT project situational management under Scrum conditions is the essence of the scientific novelty of this study, which will contribute to increasing the efficiency of IT project implementation and can be used in the relevant management system.

The purpose of the original work is to devise a model of situational IT project management in Scrum conditions based on the graph-analytical method. This would provide an opportunity to make effective decisions, taking into account the influence of external and internal environmental factors at various stages of the implementation of the IT project.

To achieve the set aim, the following tasks have been solved:

- to investigate the factors of the external and internal environment that shape the situation;
- to build graph-analytical model for situational project management of Scrum-based project in the field of information technology.

The study materials. According to Scrum [11], the effectiveness of an IT project is ensured by a process framework that defines methods and roles.

The foundation of Scrum is Sprint [12]. This is a time interval, the duration of which is from one week to one month and during which an iteration of the IT project is being implemented. The completion of Sprint is the obtaining a new working version of the IT product. Every iteration of Sprint starts with planning. Sprint Planning is the starting situation for evaluating the content of Product Backlog. Product Backlog is a document that lists the functional requirements for a product and prioritizes it. Sprint Backlog is the functional requirements from Product Backlog that the Product owner selects. All functional requirements are distributed by tasks. Sprint is made up of tasks that need to be completed. In addition, for Sprint, an important point is the formulation of the goal. The goal is the motivating factor. The implementation of Sprint Backlog tasks is aimed at achieving the goal. Having a goal also helps the team to make the optimal decision when there are several alternative ways to solve the problem.

Since the situational management of an IT project is expedient to be carried out on the basis of Scrums, the basic procedure for the implementation of an IT project in the conditions of scrums is as follows (Figure 1):

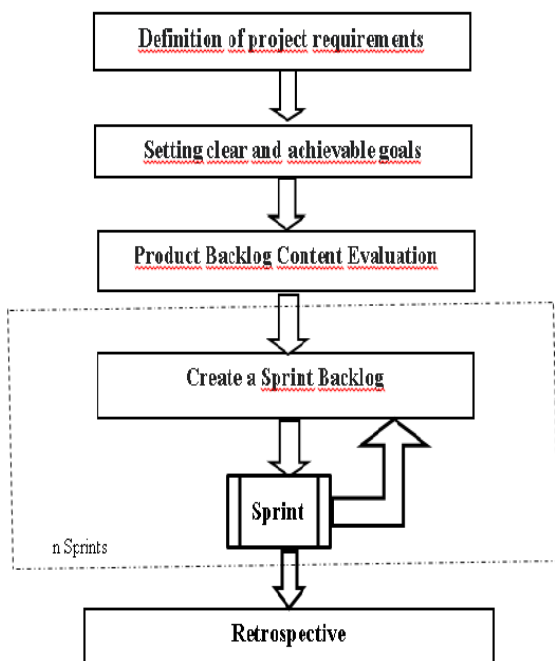


Figure 1 – Scheme for the implementation of an IT project based on the flexible Scrum methodology

During the life cycle of a project, a large number of events can occur that are difficult to foresee and study quantitatively. Therefore, it is important to establish a causal relationship between situations, situational goals, tasks, as well as the end result. Information about the factors of the external environment and the internal state of the IT project can be obtained using the method of expert assessments.

Therefore, a combined combination of intellectual and expert methods is effective. This would provide more accurate and high quality results. The reliability of the examination is ensured by a reasonable number of proposed factors, which allows not to go beyond the psychological capabilities of experts. However, the information obtained depends on the competence of experts, which is a limitation of this method [13].

Situational analysis is intended to assess the current situation during the Sprint Planning process and then during each Daily Scrum. The following procedure is proposed, which should be implemented using expert methods:

1. Examination of project situation:
 - situational goals are determined;
 - situational factors are determined;

- dynamics of influence of situational factors is being investigated;
- the situation is being formed and evaluated.

2. Development of a graph-analytical model of situational management for Sprint:

- establishing a connection between situations and situational goals;
- formation of alternative solutions aimed at achieving goals.

3. The procedure for choosing a decision aimed at achieving a situational goal.

4. Analysis of the obtained results and formation of the next situation.

To study the factors of the external and internal environment, we will apply the graph-analytical method [14, 15], which would provide the possibility of constructing a graph of the dynamics of the formation of the situation. Let the situation be characterized by internal factors f_i , $i = 1, \dots, n$ (Productivity of the IT project team, Labor intensity of the IT project, Velocity, Definition of readiness). Each factor is assigned a variable x_i , which is a function of continuous time τ and characterizes the factor:

$$x_i(\tau) = a_{i0} + a_{i1}\tau + \dots + a_{ik}\tau^k, \quad (1)$$

where a_{il} are constants, $l = 1, \dots, k$.

Over time, all situational internal factors change. Moreover, factors f_j , $j = 1, \dots, n$, $i \neq j$ characterizing the events of the external environment can influence the factors f_i . Therefore, the internal situational factor at the moment will be expressed by the formula:

$$f_i(\tau) = x_i(\tau) + \sum_j b_{ji} \times x_j(\tau), \quad (2)$$

where b_{ji} is the coefficient determining the impact of f_j on f_i . The values a_{il} , b_{ji} can be both positive and negative.

In the graph of the dynamics of the influence of situational factors, the vertices correspond to situational factors f_i , edges (f_{i1}, f_{i2}) , $i_1, i_2 = 1, \dots, n$, $i_1 \neq i_2$ correspond to the influence of f_{i1} on f_{i2} having coefficients $b_{i1,i2}$. An example of the graph of the dynamics of the influence of situational factors is given in Figure 2.

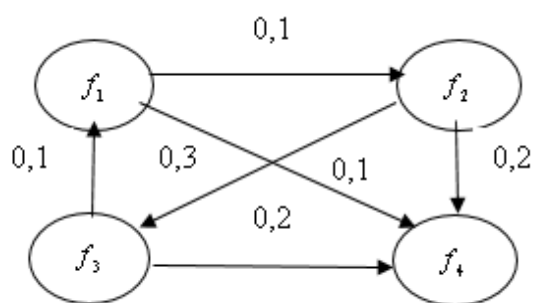


Figure 2 – Example of a graph of the dynamics of the influence of situational factors

According to (1), the functions characterizing situational factors are linear ($k = 1$). Therefore, let's express them as:

$$\begin{aligned} x_1(\tau) &= 3 + 0,4\tau \\ x_2(\tau) &= 3 - 0,4\tau \\ x_3(\tau) &= 5 + 0,2\tau \\ x_4(\tau) &= 5 - 0,2\tau \end{aligned} \quad (3)$$

Applying relations (1), (3) to the graph in Figure 2, we obtain:

$$\begin{aligned} k_1(\tau) &= 3 + 0,4\tau + 0,2(3 - 0,4\tau) = 3,6 + 0,32\tau \\ k_2(\tau) &= 3 - 0,4\tau + 0,1(3 + 0,4\tau) = 2,3 - 0,36\tau \\ k_3(\tau) &= 5 + 0,2\tau + 0,1(5 + 0,2\tau) + 0,1(5 - 0,2\tau) = \\ &= 6 + 0,2\tau \\ k_4(\tau) &= 5 - 0,2\tau + 0,1(2 + 0,5\tau) + 0,1(2 - 0,5\tau) = \\ &= 6 - 0,2\tau \end{aligned}$$

If the necessary information is available, the values of the situational factors are given by numerical variables with a certain range of values. If there is not enough relevant information, then linguistic variables with qualitative values such as "very low", "low", "medium", "high", "very high" level are entered. Qualitative values are evaluated by the number of points on a special scale (for example, a five- or ten-point scale). In this way, linguistic variables are transformed into numerical ones with a standard scope of definition. Since it is rather difficult to measure the values of the factors in the real system, we would use the single linguistic point scale presented in Table 1 for them.

The change in the value of situational factors reflects the nature of their change and corresponds to numbers from the closed interval $[0, 1]$ with the sign "+" if the indicator increases, and the sign "-" if it decreases. The stability of the

indicator is represented by the number 0. The rate of change (growth or decline) is estimated with the help of limit points on the closed interval $[0, 1]$: 0.1 – very weak; 0.3 – moderate; 0.5 – significant; 0.7 – strong; 1.0 – very strong indicator change.

Table 1 – Linguistic-scoring value of assessment of situational factors

Linguistic assessment of meaning	Point assessment
Very low	$2 > f_i \geq 0$
Low	$4 > f_i \geq 2$
Average	$6 > f_i \geq 4$
High	$9 > f_i \geq 7$
Very high	$12 > f_i \geq 10$
Super high	$f_i \geq 12$

Thus, each situation is characterized by a certain set of factors of the internal environment with the corresponding influence of factors of the external environment (Table 2).

Table 2 – Assessment of the situation through external and internal factors

factor	f_1	f_2	f_3	f_4
S_1	+0,5	+0,7	+0,3	+0,1
S_2	+0,2	+0,1	+0,2	+0,5
S_3	+0,2	+0,2	+0,2	+0,3

The graph-analytical model of situational IT project management in Scrum conditions should demonstrate cause-and-effect relationships between situations formed as a result of certain factors and decisions that should contribute to the achievement of certain goals. Depending on the obtained values of external and internal environmental factors characterizing the situation, one or another alternative solution is proposed.

The graph-analytical model of situational IT project management in Scrum conditions is a graph whose vertices correspond to situations, and arcs – to decisions aimed at achieving the set goal (Figure 3). Situations are represented by rectangles, to the upper sides of which arrows fit, according to the set goals. The root of the tree represents the starting (initial) situation.

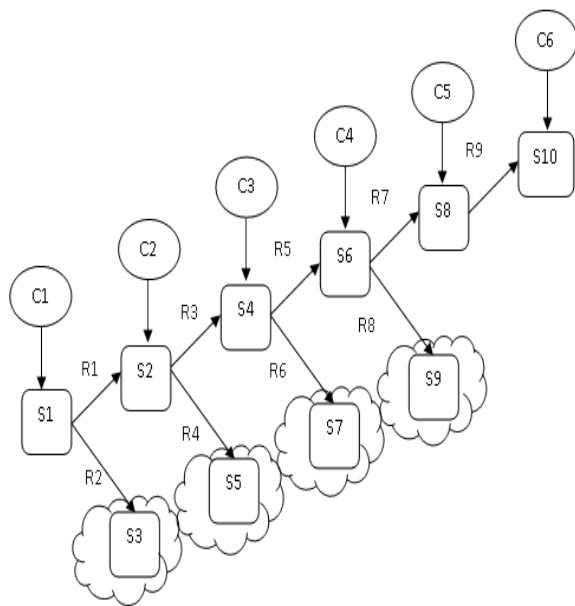


Figure 3 – Example of a graph of situations of one Sprint IT project in Scrum conditions

We would demonstrate the choice of a solution from a given alternative for a fragment of a tree using an example situation. S_1 . We form Table 3, where for each alternative solution $\{r_1\}$ we match the value of factors f_i obtained on the basis of (2). In the second column of Table 3 "weights" (priorities) of factors on a scale from 0 to 1, assessed by experts, are specified.

Table 3 – Predicted states of system indicators

factor	weight	$\{r_1\}$	$\{r_2\}$
f_1	0,28	+0,6	-0,6
f_2	0,29	+0,2	-0,2
f_3	0,20	+0,3	-0,4
f_4	0,23	+0,2	+0,1
efficiency index $W(S_1)$		0,332	-0,283

For each column of factors f_i the decision efficiency index $W(S_1)$ for the situation S_1 is calculated in the form of the sum of the values of the indicators entered in the column by their "weights":

$$W(S_1) = \sum_{i=1}^n \alpha_i f_i, \quad (4)$$

where α_i is the "weight" of the factor I_i ; n is the number of factors. Next, the graph (decision)

whose efficiency index satisfies the accepted criterion is selected. For example, if such a criterion is considered to be the maximum of the index, then a decision $\{r_1\}$ is selected on this interval and a controlling influence is formed. The procedure is repeated for the following situations.

Research results. As a result of the conducted research, the following theoretical scientific results have been obtained:

- cause-and-effect relationships between the current situation, situational goals and decisions are established. This provides opportunities to assess the situation and make an effective management decision for the IT project in Scrum conditions;

- the effects of changing situational factors on the situation, which provides real-time monitoring of the situation for the IT project under Scrum conditions, are investigated;

- an efficiency index is built for the situation in order to ensure the adoption of effective management decisions.

Discussion of results. In terms of using Scrum for IT project management, situational assessment is important both for each Daily Scrum and Sprint Planning. This will ensure the ability to adequately and accurately assess the current situation and make appropriate effective decisions that will contribute to the achievement of situational goals, as well as the avoidance of future crisis situations. And this increases the flexibility of the IT project. In such a setting, the presented scientific problem is considered for the first time.

Assessing the qualitative and quantitative impact of the factors that shape the situation provides the opportunity to study the changes in these factors that affect financial indicators after a certain period of time and does not allow to quickly respond to a change in the situation for the studied IT project during the implementation of each task. This makes it possible to make the necessary management decisions in the "real-time" mode, which allows to reduce possible costs.

The research is carried out on the basis of an academic example of project management in the field of information technologies, which provides an opportunity to illustrate situational IT project management in the context of Scrum. It is important to study the interdependent relationship between various situational factors and build an index of the situation's effectiveness. The re-

search uses the method of expert evaluations, which is a limitation of this study.

Conclusions As a result, in this paper, for the first time, the factors of the external and internal environment that shape situations are investigated. This provides opportunities to monitor changes in the situation in real time and its impact on the effectiveness of the IT project.

A graph-analytical model of situational IT project management in Scrum conditions, which provides the possibility of obtaining a cause-and-effect relationship between the exact situation, situational goals and decisions, is developed. An index of situational efficiency is proposed, which would simplify the procedure of selection and management decision-making in the process of situational management. In the perspective of further research, the obtained model can be applied in the construction of ontologies as the basis of an intelligent IT project management system. This would ensure an increase in the efficiency of the IT project due to the adoption of adequate management decisions.

References

- [1] G. I. Smith, A. L. Kolesnik, K. Lavrischeva, and O. Slabospitsky, "Improving the process of drafting families of software systems elements of agile methodologies", *Programming problems*, no. 2-3, pp. 261-270, 2010.
- [2] J. Sutherland, *Scrum: The Art of Doing Twice the Work in Half the Time* (1st ed.) Currency, 2014, p. 256. ISBN 9780385346450.
- [3] H. Kniberg, *Scrum and XP from the Trenches (Enterprise Software Development)*. C4Media, 2007, p. 140.
- [4] D. Maximini, *The Scrum Culture: Introducing Agile Methods in Organizations. Management for Professionals*. Cham: Springer, 2015, p. 26.
- [5] T. A. Prokopenko, and Ya. L. Zyelyk, "Complex method of strategic decision-making in management of technological complexes of continuous type", *Journal of Automation and Information Sciences*, vol. 49, pp. 71-79, 2017.
- [6] L. A. Zadeh, "Knowledge representation in fuzzy logic", *IEEE Trans. Knowledge and Data Eng. March*, no. 1, pp. 89-100, 1989.
- [7] A. K. Lohani, N. K. Goel, and K. K. S. Bhatia, "Comparative study of neural network, fuzzy logic and linear transfer function techniques in daily rainfall-runoff modelling under different input domains", *Hydrological Processes*, vol. 25 (2), pp. 175-193, 2011.
- [8] B. Schwenker, and T. Wulf, *Scenario-Based Strategic Planning: Developing Strategies in an Uncertain World*. Springer, 2013. ISBN: 3658028742.
- [9] Wei Dong, "Using a system dynamics simulation model to explore the validity of dynamics of Agile software development", *Graduate Theses and Dissertations*, 18486, 2021. [Online]. Available: <https://lib.dr.iastate.edu/etd/18486>.
- [10] P. Serrador, and R. Turner, "The relationship between project success and project efficiency", *Project Management Journal*, vol. 46, no. 1, pp. 30-39, 2015.
- [11] T. Prokopenko, O. Lavdanska, Y. Povolotskyi, B. Obodovskyi, and Y. Tarasenko, "Devising an integrated method for evaluating the efficiency of Scrum-based projects in the field of Information Technology", *Eastern-European Journal of Enterprise Technologies*, vol. 5, no. 3-113, pp. 46-53, 2021.
- [12] M. Andrew, "Software development roles and responsibilities in outsourcing". [Online]. Available: <https://qarea.com/blog/software-development-roles-and-responsibilities-in-outsourcing>.
- [13] S. L. Kendal, and M. Creen. *An Introduction to Knowledge Engineering*. London: Springer, 2007.
- [14] A. Ladaniuk, T. Prokopenko, and V. Reshetiuk, "The model of strategic management of organizational and technical systems, taking into account risk-based cognitive approach", *Annals of Warsaw University of Life Sciences – SGGW Agriculture (Agricultural and Forest Engineering)*, no. 63, pp. 97-104, 2014.
- [15] A. Chochowski et al., *Innovative Energy-Saving Technologies in Biotechnological Objects Control*. Kyiv: Tsentr Uchbovoi Literatury, 2014.

Т. О. Прокопенко, д-р техн. наук, професор,
e-mail: t.prokopenko@chdtu.edu.ua

О. І. Підкуйко
e-mail: o.i.pidkuiko.fitis20@chdtu.edu.ua
Черкаський державний технологічний університет
б-р Шевченка, 460, м. Черкаси, 18006, Україна

РОЗРОБКА ГРАФОАНАЛІТИЧНОЇ МОДЕЛІ СИТУАЦІЙНОГО УПРАВЛІННЯ ПРОЕКТОМ В УМОВАХ SCRUM У СФЕРІ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Запропоновано графоаналітичну модель ситуаційного управління проектом в галузі інформаційних технологій, що забезпечить вироблення ефективних управлінських рішень. При цьому враховуються всі особливості реалізації методології Scrum, що застосовується в управлінні проектами галузі інформаційних технологій. Неоднозначні ситуації, які постійно виникають у ході реалізації проєктів, впливають на їх ефективність. Тому постійний контроль ситуацій у проєкті забезпечить проєктному менеджеру можливості в оперативному прийнятті управлінських рішень, що сприяє підвищенню ефективності проєкту. Дослідження базується на застосуванні ситуаційного аналізу, що представляє комплекс технологій підготовки, прийняття та реалізації управлінського рішення на основі аналізу окремо взятої управлінської ситуації. Для кожного Sprint проєкту в умовах Scrum у ході Sprint Planning, а також у подальших Daily Scrum пропонується застосування процедури ситуаційного управління, що включає аналіз ситуаційних факторів та дослідження їх впливів, побудову графоаналітичної моделі ситуаційного управління, вибір рішення, що спрямоване на досягнення ситуативної цілі, аналіз отриманих результатів. Графоаналітична модель ситуаційного управління IT проєктом в умовах Scrum має демонструвати причинно-наслідкові зв'язки між сформованими внаслідок тих чи інших факторів ситуаціями та рішеннями, що мають сприяти досягненню визначених цілей. Встановлення причинно-наслідкових зв'язків між поточною ситуацією, ситуаційними цілями та рішеннями забезпечує можливості оцінки ситуації та прийняття ефективного управлінського рішення для IT проєкту в умовах Scrum. Дослідження впливів зміни ситуаційних факторів на ситуацію забезпечує моніторинг ситуації в режимі реального часу для IT проєкту в умовах Scrum. Запропонований індекс ефективності для ситуації дає можливість оцінки ситуації та прийняття відповідних управлінських рішень, що забезпечить підвищення ефективності проєкту.

Ключові слова: Scrum, Sprint, фактори, ситуація, граф.

УДК 004.032.26

[0000-0003-3841-7373] **Є. Є. Федоров**, д-р техн. наук, професор,
[0000-0002-6614-4133] **Т. Ю. Уткіна**, канд. техн. наук, доцент
e-mail: t.utkina@chdtu.edu.ua

Черкаський державний технологічний університет
б-р Шевченка, 460, м. Черкаси, 18006, Україна

МЕТОД КЛАСТЕРИЗАЦІЇ КВАЗІПЕРІОДИЧНОГО СИГНАЛУ НА ОСНОВІ АЛГОРИТМУ КЛОНАЛЬНОГО ВІДБОРУ

У роботі запропоновано ієрархічно-ітеративний метод кластеризації квазіперіодичного сигналу, що базується на алгоритмі клонального відбору, підвищує швидкість та точність кластеризації. Попередньо проводиться перетворення зразків (квазіперіодичних ділянок) цього сигналу до єдиного амплітудно-часового вікна на основі зсуву та масштабування за часом і амплітудою, лінійною інтерполяцією та дискретизацією за часом.

У результаті проведеного аналізу сучасних методів кластеризації квазіперіодичних сигналів встановлено, що більшість із них має один або кілька з таких недоліків: невідома точна кількість кластерів; чутливість до початкових значень центроїдів кластерів; низька ймовірність кластеризації; низька швидкість кластеризації; порівняння ділянок сигналу, що мають лише однаковий розмір; порівняння лише бінарних сигналів. Тому актуальною є розробка методу перетворення квазіперіодичного сигналу й ієрархічно-ітеративного методу кластеризації на основі алгоритму клонального відбору. Це забезпечить підвищення ефективності аналізу структури квазіперіодичного сигналу при проведенні обробки цифрових даних в інтелектуальних комп'ютерних системах ідентифікації особистості, технічної і медичної діагностики, аналізу мережевого трафіку тощо.

Наведено порівняння запропонованого авторами методу й існуючих методів кластеризації, при цьому параметр клонування $\alpha=0.1$, параметр мутації $\beta=2.5$, кількість замінних антитіл $d=0.2|N|$. Кластеризація проводилася на квазіперіодичних звуках мовлення, виголошених різними дикторами. Запропоновано метод перетворення квазіперіодичного сигналу, який перетворює зразки (квазіперіодичні ділянки) цього сигналу до єдиного амплітудно-часового вікна, за допомогою зсуву й масштабування за часом і амплітудою, інтерполяцією та дискретизацією. Це дозволяє порівнювати зразки сигналу різної довжини та з різним розмахом амплітуд. Створено метод кластеризації квазіперіодичного сигналу на основі ієрархічно-ітеративного підходу й алгоритму клонального відбору, що зменшує чутливість до початкових значень центроїдів кластерів за рахунок випадкового пошуку та забезпечує адаптивне налаштування кількості кластерів за рахунок ієрархічного підходу, а також підвищує ймовірність кластеризації до 0.98.

Ключові слова: квазіперіодичний сигнал, алгоритм клонального відбору, ієрархічно-ітеративна кластеризація, перетворення сигналу, аналіз структури сигналу.

Вступ. Сьогодні актуальним завданням є розробка інтелектуальних комп'ютерних систем, призначених для ідентифікації, аналізу та синтезу, зберігання цифрових сигналів (наприклад акустичних, графічних, вібраційних, електрограмних, комунікаційних та ін.) [1-2].

Водночас сигнали, що обробляються комп'ютерними системами ідентифікації особистості, технічної та медичної діагностики, аналізу мережного трафіку, часто є квазіперіодичними. У зв'язку з цим постає проблема побудови ефективних методів аналізу структури квазіперіодичного сигналу. Одним із засобів аналізу структури квазіперіодичного

сигналу є кластеризація, що є різновидом машинного навчання без вчителя.

Як правило, під завданням кластеризації розуміють необхідність розбиття на підмножини (кластери) заданого набору векторів визначеної розмірності відповідно до наперед визначеного критерію мінімізації спотворення (критерій оптимізації). При цьому елементи одного кластера мають схожі ознаки, а елементи різних кластерів значно відрізняються один від одного. Залежно від шляхів оцінювання спотворення й розрізняють методи кластеризації, які є важливим елементом для вирішення загальної проблеми обробки цифрових даних.

Аналіз останніх джерел і публікацій.

Сучасні методи кластеризації являють собою множину підходів та алгоритмів – від інтуїтивних та евристичних до суто математичних і поділяються на [3-5]:

- центроїдні (к-середніх, нечітких с-середніх, зсуву середнього значення та ін.) або медоїдні (к-медоїдів, субтрактивна кластеризація та ін.) [6-8]. Методи к-медоїдів та субтрактивну кластеризацію варто застосовувати, коли наявний шум або випадкові викиди, а методи к-середніх і нечітких с-середніх – у випадку швидкої обробки великого набору об'єктів. Методи субтрактивної кластеризації та зсуву середнього значення варто застосовувати, коли невідома кількість кластерів;

- моделі суміші розподілів (ЕМ-алгоритм та ін.) [9]. Ці методи варто застосовувати, коли кластери мають різний розмір, а набір об'єктів усіх кластерів можна описати сумішшю щільностей розподілів (суміш розподілів Гаусса та ін.);

- щільнісні (DBSCAN, OPTICS та ін.) [8]. Ці методи використовують у випадку наявного шуму та випадкових викидів, коли кластери мають різні форму і розмір, невідома кількість кластерів. Для формування кластерів використовують околицю Мура чи Неймана. Кластер є топологічно зв'язною областю;

- дивізивні (DIANA) [8] або агломеративні (центроїдного зв'язку, Варда, одиночного зв'язку, повного зв'язку, групового середнього) [8]. Ці методи варто застосовувати, коли кількість об'єктів, що кластеризуються, невелика. Формують дендрограму (дерево кластерів) на основі міжкластерної відстані;

- коннекціоністські, для формування кластерів використовують штучну нейронну мережу (SOM [10-11], сімейство ART [12] та ін.);

- метаввристичні [13-15], для формування кластерів використовують метавристику (генетичний алгоритм та ін.).

Існуючі методи кластеризації мають один або декілька з таких недоліків:

- невідома точна кількість кластерів;
- чутливість до початкових значень центроїдів кластерів;
- низька ймовірність кластеризації;
- низька швидкість кластеризації;
- порівняння ділянок сигналу, що мають лише однаковий розмір;
- порівняння лише бінарних сигналів.

З метою усунення більшості з них для аналізу структури квазіперіодичного сигналу пропонується виконати його кластеризацію шляхом застосування алгоритму клонального відбору [16-18]. Цей алгоритм (рисунок 1) належить до особливого класу штучних імунних систем та моделює роботу антитіл імунної системи щодо протидії проникаючому антигену завдяки вивченню його особливостей і адаптації до них.

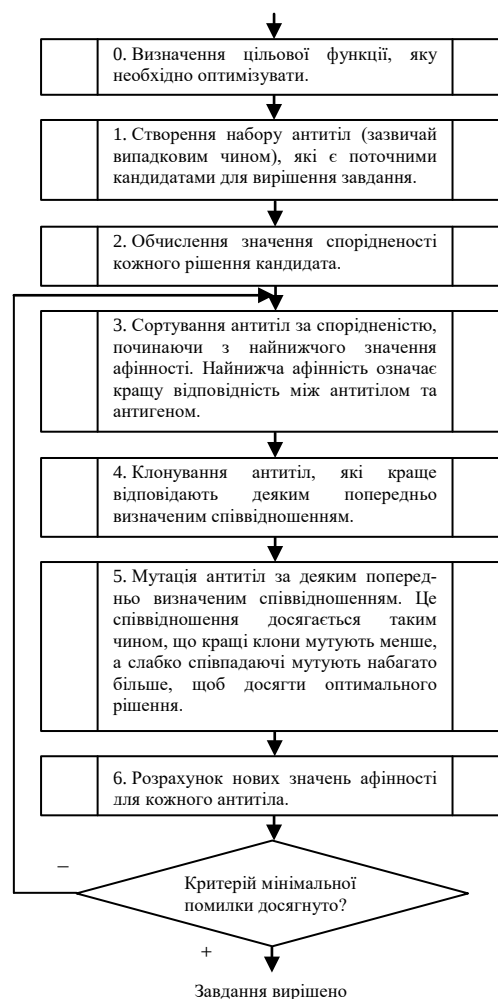


Рисунок 1 – Основні етапи алгоритму клонального відбору

Він базується на алгоритмі CLONALG, вперше запропонованому Де Кастро і Ван Зубеном [19].

На відміну від генетичного алгоритму [20], алгоритм клонального відбору використовує заміну антитіл випадково згенерованими антитілами, що підвищує різноманітність популяції та ймовірність знаходження глобального екстремуму; оператор репродукції замінений оператором клонування, а опе-

ратор схрещування не використовується, що підвищує швидкість пошуку розв'язку.

Як видно з рисунка 1, алгоритм селекції клонів починається з визначення цільової функції, яку необхідно оптимізувати. Створюються деякі можливі рішення-кандидати. Антитіла використовуються в цільовій функції для обчислення їх спорідненості. Це дозволяє визначити ті з них, які будуть клоновані для наступного кроку. Клоновані значення змінюються з попередньо визначеним співвідношенням, а спорідненість перераховується та сортується. Після проведення оцінки спорідненості та, що матиме найменше значення, і буде розв'язком, найближчим для вирішення поставленого завдання.

Мета та задачі дослідження. Метою роботи є підвищення ефективності кластеризації квазіперіодичного сигналу за рахунок синтезу методу перетворення такого сигналу й ієрархічно-ітеративного методу кластеризації на основі алгоритму клонального відбору.

Для досягнення поставленої мети необхідно вирішити такі завдання:

- розробити метод перетворення квазіперіодичного сигналу;
- запропонувати критерій якості кластеризації квазіперіодичного сигналу;
- створити ієрархічно-ітеративний метод кластеризації квазіперіодичного сигналу на основі алгоритму клонального відбору;
- провести чисельне дослідження запропонованого методу кластеризації.

1. Метод перетворення квазіперіодичного одновимірної сигналу. Включає наступні етапи:

- розбиття квазіперіодичного сигналу на дискретні зразки;
- зсув дискретних зразків за часом та амплітудою;
- інтерполяції дискретних зразків;
- зсув та масштабування неперервних зразків за часом;
- зсув та масштабування неперервних зразків за амплітудою;
- дискретизація неперервних зразків.

1.1. Розбиття квазіперіодичного сигналу на дискретні зразки. Визначимо кінцеву множину дискретних зразків квазіперіодичного сигналу, що описуються сімейством цілих обмежених фінітних дискретних функцій $X = \{x_i \mid i \in \{1, \dots, I\}\}$, у вигляді

$$x_i(n) = \begin{cases} f(n), & n \in \{N_i^{\min}, \dots, N_i^{\max}\} \\ 0, & n \notin \{N_i^{\min}, \dots, N_i^{\max}\} \end{cases}, \quad i \in \{1, \dots, I\},$$

$$A_i^{\min} = \min_n f(n), \quad n \in \{N_i^{\min}, \dots, N_i^{\max}\}, \quad i \in \{1, \dots, I\},$$

$$A_i^{\max} = \max_n f(n), \quad n \in \{N_i^{\min}, \dots, N_i^{\max}\}, \quad i \in \{1, \dots, I\},$$

де $A_i^{\min}, A_i^{\max}$ – мінімальне та максимальне значення функції x_i на відрізку $\{N_i^{\min}, \dots, N_i^{\max}\}$;

I – потужність множини зразків квазіперіодичного сигналу.

1.2. Зсув дискретних зразків за часом та амплітудою. Визначимо кінцеву множину зсунутих за часом та амплітудою дискретних зразків, що описуються кінцевим сімейством цілочисельних обмежених фінітних дискретних функцій $X^s = \{x_i^s \mid i \in \{1, \dots, I\}\}$, у вигляді

$$x_i^s(n) = \begin{cases} x_i(n + N_i^{\min}) - A_i^{\min}, & n \in \{0, \dots, N_i\} \\ 0, & n \notin \{0, \dots, N_i\} \end{cases},$$

$$i \in \{1, \dots, I\},$$

$$N_i = N_i^{\max} - N_i^{\min},$$

$$A_i = A_i^{\max} - A_i^{\min}.$$

1.3. Інтерполяції дискретних зразків.

У статті було обрано лінійну інтерполяцію, яка вимагає найменшої обчислювальної складності.

Визначимо кінцеву множину неперервних зразків, які отримані в результаті лінійної інтерполяції й описуються кінцевим сімейством дійсних обмежених фінітних неперервних функцій $\Psi = \{\psi_i \mid i \in \{1, \dots, I\}\}$, у вигляді

$$\forall t \in [-\Delta t, T_i + \Delta t], \quad \psi_i(t) =$$

$$= \sum_{n=-1}^{N_i} \chi_{(t_n, t_{n+1})}(t) \left(x_i^s(n) + \frac{x_i^s(n+1) - x_i^s(n)}{\Delta t} (t - t_n) \right) +$$

$$+ \sum_{n=-1}^{N_i+1} \chi_{\{t_n\}}(t) x_i^s(n), \quad i \in \{1, \dots, I\},$$

$$\forall t \notin [-\Delta t, T_i + \Delta t] \quad \psi_i(t) = 0,$$

$$T_i = N_i \Delta t,$$

$$t_n = n \Delta t,$$

$$\chi_B(t) = \begin{cases} 1, & t \in B; \\ 0, & t \notin B, \end{cases} \quad \text{– індикаторна функція;}$$

де Δt – крок квантування за часом.

1.4. Зсув і масштабування неперервних зразків за часом. Визначимо кінцеву множину зсунутих та масштабованих за часом неперервних зразків, що описуються кінцевим сімейством дійсних обмежених фінітних неперервних функцій $\Psi^s = \{\psi_i^s \mid i \in \{1, \dots, I\}\}$, у вигляді

$$\forall t \in [\tilde{T}^{\min}, \tilde{T}^{\max}], \psi_i^s(t) = \psi_i \left(T_i \frac{t - \tilde{T}^{\min}}{\tilde{T}^{\max} - \tilde{T}^{\min}} \right),$$

$$\forall t \notin [\tilde{T}^{\min}, \tilde{T}^{\max}], \psi_i^s(t) = 0,$$

де $[\tilde{T}^{\min}, \tilde{T}^{\max}]$ – заданий та єдиний для всіх зразків відрізок.

У статті пропонується визначити $\tilde{T}^{\min}, \tilde{T}^{\max}$ наступним чином:

$$\tilde{T}^{\min} = \Delta t,$$

$$\tilde{T}^{\max} = \left(1 + \frac{1}{I} \sum_{i=1}^I (N_i^{\max} - N_i^{\min}) \right) \Delta t.$$

1.5. Зсув і масштабування неперервних зразків за амплітудою. Визначимо кінцеву множину зсунутих та масштабованих за амплітудою неперервних зразків, що описуються кінцевим сімейством дійсних обмежених фінітних неперервних функцій $\Psi^{ss} = \{\psi_i^{ss} \mid i \in \{1, \dots, I\}\}$, у вигляді

$$\forall t \in [\tilde{T}^{\min}, \tilde{T}^{\max}],$$

$$\psi_i^{ss}(t) = \tilde{A}^{\min} + \frac{\tilde{A}^{\max} - \tilde{A}^{\min}}{\tilde{A}_i^{\max} - \tilde{A}_i^{\min}} \psi_i^s(t),$$

$$\forall t \notin [\tilde{T}^{\min}, \tilde{T}^{\max}], \psi_i^{ss}(t) = 0,$$

$$\tilde{A}_i^{\max} = \max_t \psi_i^s(t), \tilde{A}_i^{\min} = \min_t \psi_i^s(t),$$

$$t \in [\tilde{T}^{\min}, \tilde{T}^{\max}],$$

де $\tilde{A}_i^{\min}, \tilde{A}_i^{\max}$ – мінімальне та максимальне значення функції ψ_i^s на відрізку $[\tilde{T}^{\min}, \tilde{T}^{\max}]$;

$\tilde{A}^{\min}, \tilde{A}^{\max}$ – задане та єдине для всіх зразків мінімальне та максимальне значення на заданому відрізку $[\tilde{T}^{\min}, \tilde{T}^{\max}]$.

У статті пропонується визначити $\tilde{A}^{\min}, \tilde{A}^{\max}$ наступним чином:

$$\tilde{A}^{\min} = 0, \tilde{A}^{\max} = \frac{1}{I} \sum_{i=1}^I (A_i^{\max} - A_i^{\min}).$$

1.6. Дискретизація неперервних зразків. Визначимо кінцеву множину дискретних зразків, які отримані з неперервних шляхом

дискретизації за часом й описуються кінцевим сімейством цілочисельних обмежених фінітних дискретних функцій $S = \{s_i \mid i \in \{1, \dots, I\}\}$, у вигляді

$$s_i(n) = \text{round}(\psi_i^{ss}(n\Delta t)), n \in \{\tilde{N}^{\min}, \dots, \tilde{N}^{\max}\},$$

$$\tilde{N}^{\min} = \tilde{T}^{\min} / \Delta t,$$

$$\tilde{N}^{\max} = \tilde{T}^{\max} / \Delta t,$$

де $\text{round}()$ – функція, яка округляє число до найближчого цілого.

Таким чином, запропонований метод перетворення квазіперіодичного сигналу перетворює зразки (квазіперіодичні ділянки) цього сигналу до єдиного амплітудно-часового вікна, шляхом зсуву та масштабування за часом та амплітудою, інтерполяцією та дискретизацією, що дозволяє надалі порівняти зразки сигналу різної довжини та з різним розмахом амплітуд.

2. Критерій якості кластеризації квазіперіодичного сигналу. Для оцінювання методу кластеризації в роботі запропоновано критерій якості, що визначає вибір такого вектора h , який містить номери кластерів для кожного зразка та доставляє мінімум нормованій внутрішньокластерній відстані:

$$F = \frac{\sum_{i=1}^I \sum_{k=1}^K \chi_{S_k}(s_i) \|s_i - \mu_k\|^2}{(\tilde{N}^{\max} - \tilde{N}^{\min} + 1) \tilde{A}^{\max} \sum_{i=1}^I \sum_{k=1}^K \chi_{S_k}(s_i)} \rightarrow \min_h,$$

де S_k – k -й кластер;

μ_k – центр k -го кластера;

$\chi_{S_k}(s_i)$ – індикаторна функція, яка опи-

сує належність s_i кластеру S_k .

3. Ієрархічно-ітеративний метод кластеризації на основі алгоритму клонального відбору. Розглядатимемо сімейство S , отримане в результаті застосування методу перетворення квазіперіодичного сигналу, що знаходиться в єдиному амплітудно-часовому вікні, як множину об'єктів кластеризації.

Запропонований метод кластеризації включає наступні етапи:

0. Початкова кількість кластерів $K = 2$.

1. Створюється початкова популяція $H = \{h\}$, причому кожне антитіло цієї популяції представлено у вигляді $h = (h_1, \dots, h_I)$. Кожна компонента h_i взаємно однозначно

відповідає об'єкту кластеризації та містить номер кластера, що визначається у вигляді

$$h_i = \text{round}((K-1) \text{rand}() + 1),$$

де $\text{round}()$ – функція, що округлює число до найближчого цілого;

$\text{rand}()$ – функція, що генерує рівномірно розподілене випадкове число в діапазоні $[0, 1]$.

2. Для кожного антитіла h обчислюються індикаторні функції $\chi_{S_k}(s_i)$, які описують належність s_i кластеру S_k , $S_k \subseteq S$ (набувають 1, якщо належать, і 0 в іншому випадку):

$$\chi_{S_k}(s_i) = \begin{cases} 1, & h_i = k \\ 0, & h_i \neq k \end{cases}, i \in \{1, \dots, I\}, k \in \{1, \dots, K\},$$

при цьому повинні виконуватись умови:

$$\sum_{k=1}^K \chi_{S_k}(s_i) = 1, i \in \{1, \dots, I\},$$

$$\sum_{i=1}^I \chi_{S_k}(s_i) > 0, k \in \{1, \dots, K\},$$

$$\chi_{S_k}(s_i) \in \{0, 1\}, k \in \{1, \dots, K\}, i \in \{1, \dots, I\}.$$

3. Присвоюється номер ітерації $\tau = 1$.

4. Для кожного антитіла h обчислюються центроїди кластерів:

$$\mu_k(m) = \frac{\sum_{i=1}^I \chi_{S_k}(s_i) s_r(m)}{\sum_{i=1}^I \chi_{S_k}(s_i)}, k \in \{1, \dots, K\},$$

$$m \in \{\tilde{N}^{\min}, \dots, \tilde{N}^{\max}\}.$$

5. Для кожного антитіла h обчислюється значення фітнес-функції.

Фітнес-функція (цільова функція) визначена у вигляді

$$F(h) = \frac{\sum_{i=1}^I \sum_{k=1}^K \chi_{S_k}(s_i) \|s_i - \mu_k\|^2}{(\tilde{N}^{\max} - \tilde{N}^{\min} + 1) \tilde{A}^{\max} \sum_{i=1}^I \sum_{k=1}^K \chi_{S_k}(s_i)} \rightarrow \min_h,$$

$$\|s_i - \mu_k\| = \sqrt{\sum_{m=\tilde{N}^{\min}}^{\tilde{N}^{\max}} |s_i(m) - \mu_k(m)|}, i \in \{1, \dots, I\},$$

$$k \in \{1, \dots, K\}.$$

$$\hat{b}_q = \begin{cases} 1, & (\text{rand}() \geq p(h) \wedge (b_q = 1)) \vee ((\text{rand}() < p(h)) \wedge (b_q = 0)); \\ 0, & (\text{rand}() \geq p(h) \wedge (b_q = 0)) \vee ((\text{rand}() < p(h)) \wedge (b_q = 1)), \end{cases}$$

$$q \in \{1, \dots, I \cdot E\},$$

6. Для кожного антитіла h обчислюється афінність.

Афінність визначає близькість поточного антитіла до кращого антитіла та обчислюється на основі функції корисності у вигляді

$$\Phi(h) = 1 - \frac{F(h) - \min_{h \in H} F(h)}{\max_{h \in H} F(h) - \min_{h \in H} F(h)},$$

де $\Phi(h) \in [0, 1]$.

Якщо $\Phi(h) = 1$, то i -те антитіло є кращим. Якщо $\Phi(h) = 0$, то i -те антитіло є гіршим.

7. Для кожного антитіла h виконується оператор клонування.

Оператор клонування грає роль, аналогічну до оператора репродукції генетичного алгоритму.

Зазвичай кількість клонів q для кожного антитіла h визначається як

$$q = \text{round}(\alpha | H|),$$

де α – параметр клонування, $\alpha > 0$.

У результаті застосування оператора клонування формується множина клонів антитіла $Z = \{z\}$, причому кожний клон представлений у вигляді $z = (z_1, \dots, z_I)$.

8. Для кожного клону z виконується оператор мутації.

Оператор мутації дозволяє отримати з клонів антитіла нові антитіла з властивостями, що різко відрізняються за Евклідовою метрикою.

Перед мутацією кожний цілочисельний клон $z = (z_1, \dots, z_I)$ взаємно однозначно відображається в бінарний клон виду $b = (b_1, \dots, b_{I \cdot E})$ у вигляді

$$b_{iE+1-e} = \lceil 2^{1-e} \cdot z_i \rceil \bmod 2, i \in \overline{1, I}, e \in \overline{1, E},$$

де E – кількість біт для кодування однієї компоненти клону;

$\lceil \cdot \rceil$ – функція, що визначає цілу частину.

Мутація над кожним бінарним клоном b виконується у вигляді

$p(h) = e^{-\beta \Phi(h)}$ – ймовірність мутації антитіла h ,

де β – параметр мутації, $\beta > 0$. Чим більше β , тим менша ймовірність мутації.

Після завершення мутації виконується взаємно однозначне відображення кожного бінарного клону $\hat{b} = (\hat{b}_1, \dots, \hat{b}_{I-E})$ в цілочисельний клон $\hat{z} = (\hat{z}_1, \dots, \hat{z}_I)$ у вигляді

$$\hat{z}_i = \sum_{e=1}^E (2^{E-e} \cdot \hat{b}_{iE+1-e}), \quad i \in \overline{1, I}.$$

У результаті застосування оператора мутації формується множина клонів $\hat{Z} = \{\hat{z}\}$, що мутували.

9. Виконується оператор редукції.

Як оператор редукції використовується схема селекції, яка забезпечує спрямованість пошуку (найкращі антитіла зберігаються) і полягає в наступному. Антитіла попередньої популяції і клони, що мутували, об'єднуються в проміжну популяцію й впорядковуються за афінністю. У нову популяцію відбираються $|H|$ перших (найкращих за афінністю) особин проміжної популяції.

10. Виконується оператор заміни.

Для ширшого дослідження простору пошуку використовується заміна останніх (найгірших за афінністю) антитіл популяції новими антитілами. Кількість антитіл d , що замінюються, фіксована і становить від 5 % до 20 % розміру популяції.

Кожне нове антитіло представлене у вигляді $h = (h_1, \dots, h_I)$, а кожна його компонента генерується у вигляді

$$h_i = \text{round}((K-1)\text{rand}() + 1), \quad i \in \overline{1, I}.$$

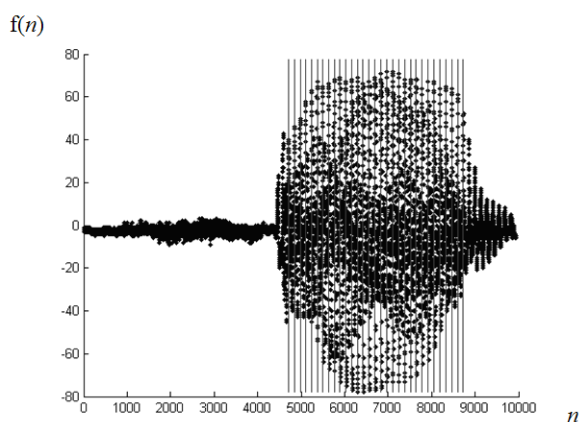


Рисунок 2 – Розбиття квазіперіодичного звуку «о» слова «со» на дискретні зразки

11. Визначається найкраще антитіло за фітнес-функцією

$$h^* = \arg \min_{h \in H} F(h).$$

12. Перевіряється умова завершення.

При $\frac{F(h^*)}{\bar{A}^{\max} \sqrt{\bar{N}^{\max}}} \leq \varepsilon$ отримано задовіль-

ний результат методу кластеризації h^* , де ε – задане значення точності кластеризації.

Якщо $\frac{F(h^*)}{\bar{A}^{\max} \sqrt{\bar{N}^{\max}}} > \varepsilon$ і $\tau < \tau^{\max}$, то

$\tau = \tau + 1$, перехід на 4.

Якщо $\frac{F(h^*)}{\bar{A}^{\max} \sqrt{\bar{N}^{\max}}} > \varepsilon$ і $\tau = \tau^{\max}$, то

$K = K + 1$, перехід на 1.

У результаті роботи методу кластеризації буде визначено структуру квазіперіодичного сигналу, що задається h^* .

Відтак, запропонований метод кластеризації квазіперіодичного сигналу на основі ієрархічно-ітеративного підходу та алгоритму клонального відбору зменшує чутливість до початкових значень центроїдів кластерів за рахунок випадкового пошуку й забезпечує адаптивне налаштування кількості кластерів завдяки ієрархічному підходу.

Результати досліджень. Чисельне дослідження проводилося на основі квазіперіодичного звуку мовленнєвого сигналу на комп'ютері з процесором Intel Pentium Quad-Core з тактовою частотою 2.58 ГГц в середовищі MATLAB.

На рисунках 2-7 зображено приклади роботи методу перетворення квазіперіодичного одновимірної сигналу.

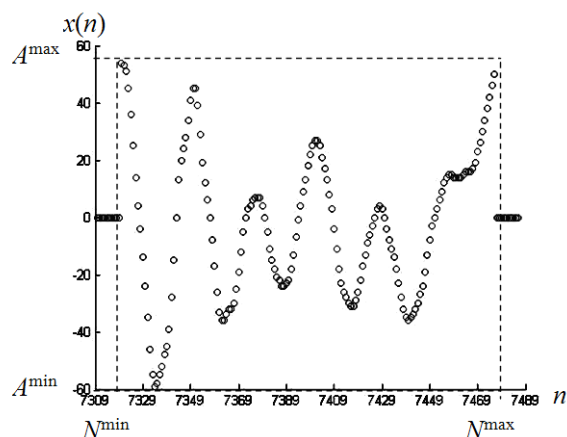


Рисунок 3 – Початковий дискретний зразок

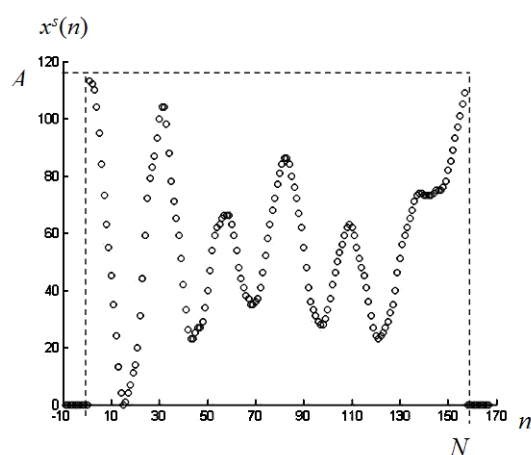


Рисунок 4 – Зсунутий за часом та амплітудою дискретний зразок

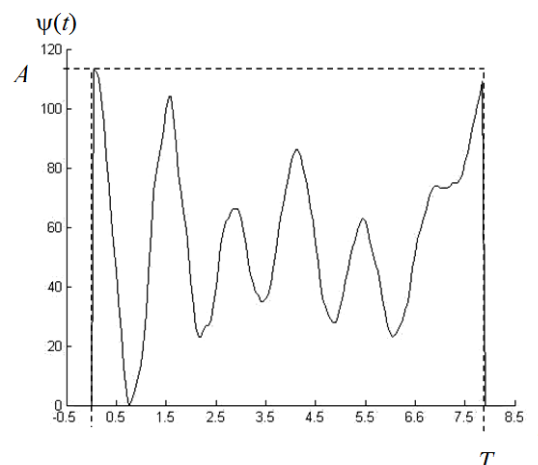


Рисунок 5 – Неперервний зразок

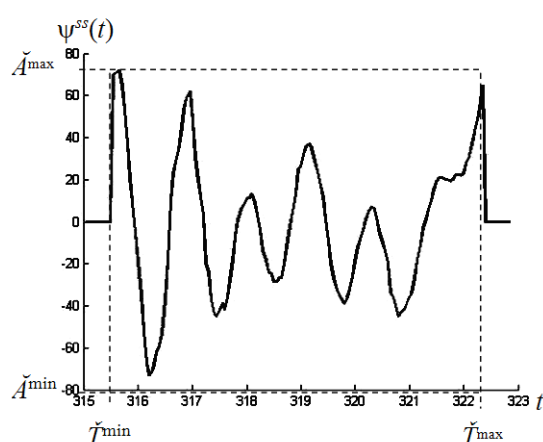


Рисунок 6 – Зсунутий і масштабований за часом та амплітудою неперервний зразок

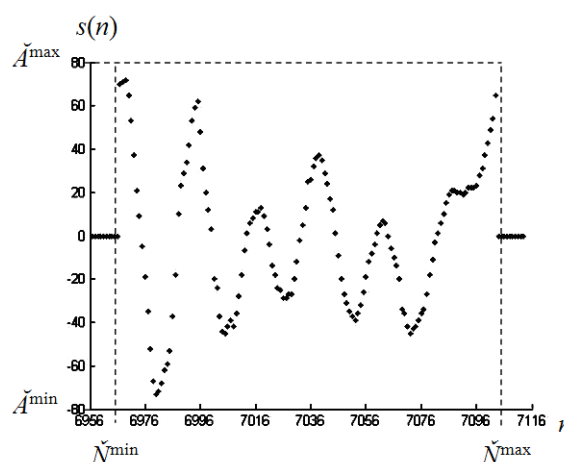


Рисунок 7 – Результуючий дискретний зразок

На рисунку 2 зображено розбиття квазі-періодичного звуку «о» слова «со» на дискретні зразки. Наведений ансамбль підтверджує складність візуального сприйняття акустичної характеристики зразка. Тому для усунення цього недоліку з дискретних зразків вибирається один, який і буде підданий перетворенню.

На рисунку 3 зображено початковий дискретний зразок, що знаходиться на відрізку $\{N_i^{\min}, \dots, N_i^{\max}\}$ з мінімальним та максимальним значенням амплітуди $A_i^{\min}, A_i^{\max}$ і піддається подальшим перетворенням, показаним на рисунках 3-6.

На рисунку 4 зображено дискретний зразок після зсуву за часом та амплітудою, що знаходиться на відрізку $\{0, \dots, N\}$ з мінімальним та максимальним значенням $0, A$.

На рисунку 5 зображено неперервний зразок сигналу, отриманий з дискретного зразка в результаті лінійної інтерполяції з частотою 22050 Гц та кроком дискретизації за часом $\Delta t = 1/22050$, що знаходиться на відрізку $[0, T]$ з мінімальним та максимальним значенням $0, A$.

На рисунку 6 зображено неперервний зразок сигналу після зсуву та масштабування за часом та амплітудою, що знаходиться на відрізку $[\tilde{T}^{\min}, \tilde{T}^{\max}]$ з мінімальним та максимальним значенням $\tilde{A}^{\min}, \tilde{A}^{\max}$.

На рисунку 7 зображено результуючий дискретний зразок сигналу, отриманий з неперервного зразка в результаті дискретизації з частотою 22050 Гц та кроком дискретизації $\Delta t = 1/22050$. Перетворений зразок знаходиться на відрізку $\{\tilde{N}^{\min}, \dots, \tilde{N}^{\max}\}$ з мінімальним та максимальним значенням $\tilde{A}^{\min}, \tilde{A}^{\max}$.

У таблиці 1 наведено порівняння запропонованого авторами методу та існуючих методів.

тодів кластеризації, при цьому параметр клонування $\alpha = 0.1$, параметр мутації $\beta = 2.5$, кількість антитіл $d = 0.2 |H|$, що замінюються.

Таблиця 1 – Оцінка методів кластеризації

Метод	Ймовірність кластеризації (%)
центроїдний (к-середніх)	0.80
центроїдний (нечітких с-середніх)	0.83
модель суміші розподілів (ЕМ)	0.85
коннекціоністський (мережа SOM)	0.82
медоїдний (субтрактивна кластеризація)	0.90
коннекціоністський (мережа ART)	0.93
авторський	0.98

Під ймовірністю кластеризації малося на увазі відношення правильно класифікованих зразків квазіперіодичного сигналу до їх загальної кількості. Кластеризація проводилася на квазіперіодичних звуках мовлення, вимовлених різними дикторами.

Обговорення результатів. Результат, наведений у таблиці 1, показує, що авторський метод забезпечує високу ймовірність кластеризації (використовує попередню обробку та адаптивне налаштування кількості кластерів).

На основі проведеного експерименту можна зробити висновок, що центроїдні методи кластеризації (к-середніх, нечітких с-середніх), метод моделі суміші розподілів (ЕМ) і коннекціоністський метод (мережа SOM) мають найменшу ймовірність кластеризації, оскільки мають фіксовану кількість кластерів.

Коннекціоністський метод (мережа ART) та медоїдний метод (субтрактивна кластеризація) мають більш високу ймовірність кластеризації, оскільки мають адаптивне налаштування кількості кластерів.

Візуалізація результатів (див. рисунки 2-7) застосування методу перетворення квазіперіодичного сигналу і чисельного дослідження ієрархічно-ітеративного методу кластеризації квазіперіодичного сигналу на основі алгоритму клонального відбору підтверджує ефективність запропонованого

методу для аналізу структури квазіперіодичного сигналу.

Висновки. У статті в межах створення методу перетворення та кластеризації квазіперіодичного сигналу було вирішено такі завдання:

– розроблено метод перетворення квазіперіодичного сигналу;

– створено ієрархічно-ітеративний метод кластеризації квазіперіодичного сигналу на основі алгоритму клонального відбору;

– проведено чисельне дослідження запропонованого методу кластеризації.

Наукова новизна проведеного дослідження полягає в наступному:

1. Удосконалено метод перетворення квазіперіодичного сигналу, який перетворює зразки (квазіперіодичні ділянки) цього сигналу до єдиного амплітудно-часового вікна, за допомогою зсуву та масштабування за часом і амплітудою, інтерполяції та дискретизації, що дозволяє надалі порівнювати зразки сигналу різної довжини і з різним розмахом амплітуд.

2. Отримав подальший розвиток метод кластеризації квазіперіодичного сигналу на основі ієрархічно-ітеративного підходу та алгоритму клонального відбору, який зменшує чутливість до початкових значень центроїдів кластерів за рахунок випадкового пошуку та забезпечує адаптивне налаштування кількості кластерів за рахунок ієрархічного підходу. Це підвищує можливість кластеризації до 0.98.

3. Отримав подальший розвиток критерій якості кластеризації квазіперіодичного сигналу, що дозволяє отримати нормовану внутрішньокластерну відстань, що, в свою чергу, дозволяє встановити єдиний поріг для оцінювання якості кластеризації.

Практичне значення. Запропонований метод дозволяє ставити і вирішувати завдання перетворення та кластеризації мовленнєвого або діагностичного сигналу, а також мережевого трафіку, що використовується для аналізу та зберігання біометричної, діагностичної та мережевої інформації в інтелектуальних комп'ютерних системах.

Перспективи подальших досліджень. Надалі передбачається досліджувати питання впливу методів перетворення та кластеризації на виявлення та класифікацію періодів у одновимірних квазіперіодичних сигналах.

Список використаних джерел

- [1] K. S. Thyagarajan, *Introduction to Digital Signal Processing using MATLAB with Application to Digital Communications*. Cham, Switzerland: Springer, 2020.
- [2] P. R. Hill, *Audio and Speech Processing with MATLAB*. Boca Raton, FL, USA: CRC Press, 2019.
- [3] D. Barber, *Bayesian Reasoning and Machine Learning*. Cambridge, GB: Cambridge University Press, 2013.
- [4] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning with Applications in R*. New York, USA: Springer, 2013.
doi: 10.1007/978-1-4614-7138-7.
- [5] T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical Learning with Sparsity. The Lasso and Generalizations*. New York, USA: Chapman and Hall/CRC, 2015.
doi: 10.1201/b18401.
- [6] А. Н. Ткаченко, О. Ф. Грийо Тукало, О. В. Дзись, и С. М. Лаховец, "Метод кластеризации на основе последовательного запуска k-средних с усовершенствованным выбором кандидата на новую позицию вставки", *Наукові праці ВНТУ*, № 2, с. 1-10, 2012.
- [7] M. J. Brusco, E. Shireman, and D. Steinley, "A comparison of latent class, k-means, and k-median methods for clustering dichotomous data", *Psychological Methods*, 22 (3), pp. 563-580, 2017.
doi: 10.1037/met0000095.
- [8] C. C. Aggarwal, and C. K. Reddy, *Data Clustering: Algorithms and Applications*. Boca Raton, FL, USA: CRC Press, 2014.
- [9] Z. Fu, and L. Wang, "Color image segmentation using Gaussian mixture model and EM algorithm", in *Proc. Multimedia and Signal Processing. Communications in Computer and Information Science (CMSP'2012)*, 2012, vol. 346, pp. 61-66.
doi: 10.1007/978-3-642-35286-7_9.
- [10] T. Kohonen, *Self-Organizing and Associative Memory*, 3rd ed. Berlin, Heidelberg, Germany: Springer, 2012.
doi: 10.1007/978-3-642-88163-3.
- [11] T. Kohonen, "Essentials of the self-organizing map", *Neural Networks: The Official Journal of the Inter-National Neural Network Society*, vol. 37, pp. 52-65, 2013.
doi: 10.1016/j.neunet.2012.09.018.
- [12] K.-L. Du, and M. N. S. Swamy, *Neural Networks and Statistical Learning*. London, GB: Springer-Verlag, 2014.
doi: 10.1007/978-1-4471-5571-3.
- [13] Ю. А. Скобцов, и Е. Е. Федоров, *Метаэвристики*. Донецк, Украина: Ноулидж (Донецкое отд.), 2013.
- [14] J. Radosavljević, *Metaheuristic Optimization in Power Engineering*. New York, USA: Institution of Engineering and Technology, 2018.
doi: 10.1049/pbpo131e.
- [15] E. Alba, A. Nakib, and P. Siarry, *Metaheuristics for Dynamic Optimization*. Berlin, Germany: Springer-Verlag, 2013.
doi: 10.1007/978-3-642-30665-5.
- [16] K.-L. Du, and M. N. S. Swamy, *Search and Optimization by Metaheuristics. Techniques and Algorithms Inspired by Nature*. Charm, Switzerland: Springer, 2016.
doi: 10.1007/978-3-319-41192-7.
- [17] O. Bozorg Haddad, M. Solgi, and H. Loaiciga, *Meta-heuristic and Evolutionary Algorithms for Engineering Optimization*. Hoboken, New Jersey, USA: Wiley & Sons, 2017.
doi: 10.1002/9781119387053.
- [18] X.-S. Yang, *Nature-Inspired Algorithms and Applied Optimization*. Charm, Switzerland: Springer, 2018.
doi: 10.1007/978-3-319-67669-2.
- [19] L. N. De Castro, and F. J. Von Zuben, "Learning and optimization using the clonal selection principle", *IEEE Transactions on Evolutionary Computation, Special Issue on Artificial Immune Systems*, vol. 6, no. 3, pp. 239-251, 2002.
- [20] E. D. Ülker, and S. Ülker, "Comparison study for clonal selection algorithm and genetic algorithm", *International Journal of Computer Science & Information Technology (IJCSIT)*, vol. 4, no. 4, pp. 107-118, 2012.
doi: 10.5121/ijcsit.2012.4410

References

- [1] K. S. Thyagarajan, *Introduction to Digital Signal Processing using MATLAB with Application to Digital Communications*. Cham, Switzerland: Springer, 2020.
- [2] P. R. Hill, *Audio and Speech Processing with MATLAB*. Boca Raton, FL, USA: CRC Press, 2019.

- [3] D. Barber, *Bayesian Reasoning and Machine Learning*. Cambridge, GB: Cambridge University Press, 2013.
- [4] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning with Applications in R*. New York, USA: Springer, 2013.
doi: 10.1007/978-1-4614-7138-7.
- [5] T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical Learning with Sparsity. The Lasso and Generalizations*. New York, USA: Chapman and Hall/CRC, 2015.
doi: 10.1201/b18401.
- [6] A. N. Tkachenko, O. F. Griyo Tukalo, O. V. Dzisy, and S. M. Lakhovets, "Clustering method based on sequential launch of k-means with improved selection of a candidate for a new insertion position", *Naukovi pratsi VNTU*, no. 2, pp. 1-10, 2012 [in Russian].
- [7] M. J. Brusco, E. Shireman, and D. Steinley, "A comparison of latent class, k-means, and k-median methods for clustering dichotomous data", *Psychological Methods*, 22 (3), pp. 563-580, 2017.
doi: 10.1037/met0000095.
- [8] C. C. Aggarwal, and C. K. Reddy, *Data Clustering: Algorithms and Applications*. Boca Raton, FL, USA: CRC Press, 2014.
- [9] Z. Fu, and L. Wang, "Color image segmentation using Gaussian mixture model and EM algorithm", in *Proc. Multimedia and Signal Processing. Communications in Computer and Information Science (CMSP 2012)*, 2012, vol. 346, pp. 61-66.
doi: 10.1007/978-3-642-35286-7_9
- [10] T. Kohonen, *Self-Organizing and Associative Memory*, 3rd ed. Berlin, Heidelberg, Germany: Springer, 2012.
doi: 10.1007/978-3-642-88163-3
- [11] T. Kohonen, "Essentials of the self-organizing map", *Neural Networks: The Official Journal of the Inter-National Neural Network Society*, vol. 37, pp. 52-65, 2013.
doi: 10.1016/j.neunet.2012.09.018.
- [12] K.-L. Du, and M. N. S. Swamy, *Neural Networks and Statistical Learning*. London, GB: Springer-Verlag, 2014.
doi: 10.1007/978-1-4471-5571-3.
- [13] Yu. A. Skobtsov, and E. E. Fedorov, *Metaheuristics*. Donetsk, Ukraine: Noulidzh (Donetsk branch), 2013 [in Russian].
- [14] J. Radosavljević, *Metaheuristic Optimization in Power Engineering*. New York, USA: Institution of Engineering and Technology, 2018.
doi: 10.1049/pbpo131e.
- [15] E. Alba, A. Nakib, and P. Siarry, *Metaheuristics for Dynamic Optimization*. Berlin, Germany: Springer-Verlag, 2013.
doi: 10.1007/978-3-642-30665-5.
- [16] K.-L. Du, and M. N. S. Swamy, *Search and Optimization by Metaheuristics. Techniques and Algorithms Inspired by Nature*. Charm, Switzerland: Springer, 2016.
doi: 10.1007/978-3-319-41192-7.
- [17] O. Bozorg Haddad, M. Solgi, and H. Loaiciga, *Meta-heuristic and Evolutionary Algorithms for Engineering Optimization*. Hoboken, New Jersey, USA: Wiley & Sons, 2017.
doi: 10.1002/9781119387053.
- [18] X.-S. Yang, *Nature-Inspired Algorithms and Applied Optimization*. Charm, Switzerland: Springer, 2018.
doi: 10.1007/978-3-319-67669-2.
- [19] L. N. De Castro, and F. J. Von Zuben, "Learning and optimization using the clonal selection principle", *IEEE Transactions on Evolutionary Computation, Special Issue on Artificial Immune Systems*, vol. 6, no. 3, pp. 239-251, 2002.
- [20] E. D. Ülker, and S. Ülker, "Comparison study for clonal selection algorithm and genetic algorithm", *International Journal of Computer Science & Information Technology (IJCSIT)*, vol. 4, no. 4, pp. 107-118, 2012.
doi: 10.5121/ijcsit.2012.4410.

E. E. Fedorov, Dr. Tech. Sc., Professor,
T. Yu. Utkina, Ph. D., Associate Professor
e-mail: t.utkina@chdtu.edu.ua
Cherkasy State Technological University
Shevchenko blvd, 460, Cherkasy, 18006, Ukraine

METHOD OF CLUSTERIZATION OF QUASIPERIODIC SIGNAL BASED ON CLONAL SELECTION ALGORITHM

The paper proposes a hierarchical-iterative method of clustering of a quasi-periodic signal based on clonal selection algorithm, which increases the speed and accuracy of clustering. Preliminary transformation of samples (quasi-periodic areas) of this signal into a single amplitude-time window based on shift and scaling in time and amplitude, linear interpolation and time sampling is performed.

Signals processed by intelligent computer systems for identification, analysis and synthesis, storage of digital signals (such as acoustic, graphic, vibration, electrogram, communication ones, etc.) are often quasi-periodic. This raises the problem of constructing effective methods for analyzing the structure of a quasi-periodic signal. One of the means of analyzing the structure of a quasi-periodic signal is clustering, which is a type of machine learning without a teacher.

As a result of the analysis of modern methods of clustering of quasi-periodic signals, it has been found that most of them have one or more of the following disadvantages: unknown exact number of clusters; sensitivity to initial values of the centroids of clusters; low probability of clustering; low speed of clustering; comparison of signal areas that have only the same size; comparison of only binary signals.

Therefore, it is important to develop a method for transforming a quasi-periodic signal and a hierarchical-iterative method of clustering based on a clonal selection algorithm. This will increase the efficiency for the analysis of quasi-periodic signal structure in digital data processing in intelligent computer systems of identity identification, technical and medical diagnostics, network traffic analysis, etc.

A comparison of the method proposed by the authors and existing clustering methods is given, with the cloning parameter $\alpha = 0.1$, the mutation parameter $\beta = 2.5$, the number of replacement antibodies $d = 0.2 \cdot H$. Clustering has been performed on quasi-periodic speech sounds uttered by different speakers. A method for converting a quasi-periodic signal, which converts samples (quasi-periodic areas) of this signal into a single amplitude-time window by shifting and scaling in time and amplitude, interpolation and sampling, is proposed. This allows to compare signal samples of different lengths and with different amplitudes. A method for clustering a quasi-periodic signal based on a hierarchical-iterative approach and a clonal selection algorithm has been developed, which reduces the sensitivity to initial values of cluster centroids due to random search and provides adaptive adjustment of the cluster number due to the hierarchical approach, and also increases the probability of clustering to 0.98.

Keywords: quasi-periodic signal, clonal selection algorithm, hierarchical-iterative clustering, signal transformation, signal structure analysis.

[0000-0003-2334-2276] **І. С. Скітер**¹, канд. фіз.-мат. наук, доцент,

e-mail: skiteris@ukr.net

[0000-0002-2118-4748] **М. В. Савельєв**^{1,2}, канд. техн. наук,

e-mail: m.saveliev@isnpp.kiev.ua

[0000-0002-1598-663X] **С. В. Купріяничук**¹,

e-mail: kupriianchuk.s@gmail.com

[0000-0001-8681-2840] **Д. О. Хоменко**^{1,2}

e-mail: khomenko.amid@gmail.com

¹Інститут проблем безпеки АЕС НАН України

вул. Кірова, 36-а, м. Чорнобиль, Київська область, 07270, Україна

²Інститут проблем математичних машин та систем НАН України

пр-т Академіка Глушкова, 42, м. Київ, 03187, Україна

АЛГОРИТМ МОНІТОРИНГУ ПОТУЖНОСТЕЙ КРИТЕРІЇВ ТЕСТУВАННЯ ЕКСПОНЕНЦІАЛЬНОСТІ ЩІЛЬНОСТІ ПОТОКУ НЕЙТРОНІВ В «ОБ'ЄКТІ УКРИТТЯ»

Системний моніторинг щільності потоку нейтронів від паливовмісних матеріалів в «Об'єкті Укриття» потребує постійного контролю та визначення параметрів динаміки з метою контролю ядерної безпеки. Запропоновано алгоритм ідентифікації переходів форми трендів динамічних рядів потоків нейтронів в експоненціальну залежність для фіксації подій, які можуть свідчити про зміну стану та поведінки паливовмісних матеріалів. У роботі проводиться моніторинг потужностей статистичних критеріїв, які можуть бути використані при визначенні моментів / часових інтервалів переходу динамічних рядів щільності потоку нейтронів до експоненціального закону розподілу. Запропоновано набір критеріїв, який найбільш доцільно використовувати при перевірці експоненціального закону розподілу, гамма-розподілу, розподілу Пуассона та Вейбула. Оцінка потужностей статистичних критеріїв була проведена на різних об'ємах вибірок, без урахування конкуруючих гіпотез. Встановлено потужності трьох груп критеріїв тестування експоненціальності для різної дискретизації вхідних масивів даних. Рекомендовано мотивоване використання критеріїв для різних об'ємів вибірок. Проведені дослідження є основою для проектування підсистеми тестування експоненціальності в автоматизованій системі підтримки прийняття рішень по забезпеченню та підвищенню рівня ядерної безпеки комплексу «Новий Безпечний Конфайнмент – «Об'єкт Укриття». Отримані результати разом з визначеними наборами оцінки аномальності в дискретизованих масивах даних щільності потоку нейтронів є основою для побудови алгоритмів функціонування автоматизованих систем контролю оцінки відхилень від безпеки в комплексі «Новий Безпечний Конфайнмент – «Об'єкт Укриття».

Ключові слова: ядерно-небезпечні скупчення, імітаційне моделювання, аналіз динамічних рядів, оперативна діагностика, перевірка гіпотез, статистичні тести, підтримка прийняття рішень.

Вступ. В роботі [1] зазначається, що на сьогоднішні першочерговими задачами є вирішення комплексу інформаційних завдань, пов'язаних з проблематикою оперативної діагностики технічного стану критичних елементів та систем головного обладнання АЕС. Зазначається, що розв'язок таких завдань базується на створенні математичних моделей процесів та використанні відповідних обчислювальних методів для використання діагностичної інформації, що міститься в сигналах

датчиків інформаційно значущих діагностичних параметрів станів об'єктів; розробці математичних моделей причинно-наслідкових зв'язків, динаміки зміни спектральних параметрів великих масивів поточних технологічних параметрів тощо.

Процес аналізу динамічних рядів даних параметрів щільності потоків нейтронів (ЩПН), на нашу думку, якісно може включати два паралельні процеси: визначення аномальних подій та визначення зміни станів чи

моментів переходу до таких станів. В роботі [2] запропоновано комплекс максимально ефективних з обчислювальної точки зору статистичних критеріїв та методику їх використання для аналізу аномальних значень у динамічних рядах ЩПН залежно від необхідної потужності критеріїв та об'єму вибірки, яка досліджується. Для створення програмного забезпечення, яке дозволить в автоматичному режимі проводити оцінку станів ЩПН, необхідним також є дослідження проблеми оцінки моментів / інтервалів переходу динамічних рядів, які досліджуються, до експоненціального закону розподілу та його параметрів. Саме такий перехід буде означати наявність критичного стану і можливість визначення значущих діагностичних параметрів ЩПН для кожної складової. Практична актуальність цього підходу полягає в можливості удосконалення існуючих методик та/або використання при прогнозуванні критичних станів ЩПН від потенційно ядерно-небезпечних скупчень паливовмісних матеріалів (ПВМ) у комплексі «Новий безпечний конфайнмент – «Об'єкт Укриття» («НБК–ОУ»).

Після встановлення НБК в проектне положення детектори штатної системи контролю ядерної безпеки почали реєструвати збільшення ЩПН на периферії потенційно ядерно-небезпечного скупчення ПВМ. Ці процеси відбуваються за рахунок зміни температурного і вологісного режиму. В результаті цього відбувається втрата вологи в гіпотетично ядерно-небезпечних скупченнях ПВМ та, відповідно, зростання коефіцієнта розмноження нейтронів, що ставить під сумнів визначені гарантії ядерної безпеки ПВМ комплексу «НБК–ОУ».

На сьогодні науковці ІПБ АЕС НАН України та спеціалісти ДСП ЧАЕС розробили «Програму моніторингу ПВМ ОУ», відповідно до якої розробляються не тільки методи та підходи експериментального підтвердження ядерної безпеки, але й способи ефективного оперативного контролю параметрів ПВМ. Таким чином, алгоритм аналізу експоненціальної залежності зміни ЩПН, на нашу думку, буде використаний в автоматизованих системах підтримки прийняття рішень по забезпеченню та підвищенню рівня ядерної безпеки комплексу «НБК–ОУ».

Тестування на експоненціальність – це загальний термін для тестів відповідності для експоненційних розподілів. Для дослідження експоненціальності динамічного ряду за до-

могою параметричних методів використовується велика кількість статистичних тестів.

Так, в роботі [3] розглянуто 22 статистичні критерії перевірки експоненціальності. Вони мають різне прикладне напруження, різні алгоритми обчислень та формування висновків, ступінь значущості та ін. Рекомендації щодо їх використання для різних об'ємів вибірок не систематизовані та різномірні відносно потужностей приведених критеріїв.

В роботі [4] описано 25 тестів, які частково перекриваються з тестами, описаними в [3]. У цій роботі наголошується на практичному використанні зазначених тестів для аналізу надійності, при перевірці асиметричних розподілів та аналізу даних на дискретизованих часових інтервалах. Крім того, визначено потужності критеріїв для різних рівнів значущості від $\alpha=0,01$ до $\alpha=0,1$, але для обмежених значень вибірок об'ємами 12, 20 та 28 одиниць.

В роботі [5] проводиться перевірка експоненціальності попередньо цензурованих даних за допомогою дивергенції Тсаліса, визначаються критерії якості експоненціальності на основі оцінки розходжень даних для різних рівнів цензурування. Методика, пропонується авторами, потребує значних інформаційних ресурсів, але має більшу потужність визначення експоненціальності порівняно з іншими тестами. Результати підтверджено моделюванням методом Монте-Карло.

Крім традиційних параметричних методів, для перевірки експоненціальності також використовуються непараметричні методи. Наприклад, в роботі [6] розроблено непараметричний тест «аналіз відмов», що базується на аналізі та розрахунку параметрів експоненціального розподілу. Використовується для якісного аналізу доцільності/недоцільності методів лікування, коли розподіл частот має неекспоненційний/експоненційний розподіл.

Методологія оцінки експоненціальності, розроблена авторами [7], базується на модифікації критерію експоненціальності Ліллієфорса. Авторами розроблено новий тест експоненціальності – Goodness-of-Fit Test (GOFT), в якому розглянуто аналіз суми всіх абсолютних різниць між експоненційною функцією кумулятивного розподілу (Cumulative Distribution Function – CDF) та вибірковою емпіричною функцією розподілу (Empirical Distribution Function – EDF). Запропонований тест простий для розуміння та обчислення.

Слід зазначити, що недоліком тесту є його обмежена потужність на великих вибірках даних. Проте на об'ємах вибірок $n \leq 50$ для рівнів значущості $\alpha=0,01$; $\alpha=0,05$; $\alpha=0,1$ потужність тесту є достатньою.

На сьогодні найповнішим оглядом статистичних тестів перевірки експоненціальності є дослідження Everestus O. Ossai, Mbanefo S. Madukaife & Abimibola V. Oladugba [8], в якому розглянуто більше 90 тестів на експоненціальність. Деякі з тестів є універсальними, тоді як інші суперечать деяким особливим класам рядів розподілу. Потужність 40 із цих різних тестів на експоненціальність наборів даних порівнюється за допомогою моделювання Монте-Карло. Порівняння проводяться для об'ємів вибірок $n=\{10, 25, 50, 100\}$ для різних груп розподілів відповідно до форми їхніх функцій при рівні значущості $\alpha=0,05$. Крім того, методи оцінки експоненціальності застосовуються до двох реальних наборів даних, а міра потужності використовується для порівняння тестів. Результати показують, що деякі тести, які дуже хороші в одній групі альтернативних розподілів, не є такими в іншій групі. Крім того, деякі тести демонструють відносно високу потужність над усіма групами альтернативних розподілів, які вивчалися, тоді як деякі інші мали низькі показники потужності. Результати, отримані з реальних наборів даних, повністю узгоджуються з результатами моделювання.

Мета та задачі дослідження. Таким чином, наявність великої кількості критеріїв оцінки експоненціальності визначає завдання їх практичного мотивованого вибору для аналізу реальних експериментальних даних. Це зумовлено тим, що наявна в публікаціях інформація не дозволяє однозначно віддати перевагу певному критерію чи обраній групі критеріїв. Відповідно, основними задачами статті є:

- визначення оцінок потужностей групи критеріїв, сформованої за ознаками обчислювальної оптимальності та ефективності;
- дослідження впливу об'єму вибірки на розподіл статистик критеріїв при справедливості перевірки гіпотези експоненціальності;
- впорядкування на основі оцінок потужностей та впливу об'ємів вибірок досліджуваної групи критеріїв за перевагою.

Метою статті є моніторинг статистичних критеріїв оцінки експоненціальності динамічних рядів даних ЩПН в комплексі «НБК–ОУ», визначення потужностей критеріїв

на основі моделювання, відбір максимально ефективних критеріїв для визначених об'ємів спостережень.

Виклад основного матеріалу. Якщо досліджуваний динамічний ряд ЩПН в гіпотетично ядерно-небезпечному скупченні ПВМ в комплексі «НБК–ОУ», який реєструється датчиками, стає відповідним експоненціальному закону розподілу, то це може свідчити про можливий передкритичний стан ядерної небезпечної системи. Тому тестування дасть змогу в режимі реального часу проводити оцінку параметрів розподілу і приймати ефективні та своєчасні технічні, технологічні та управлінські рішення.

Динамічні ряди даних, які фіксуються датчиками з періодичністю $t=1$ хв в режимі 24/7, являють собою масиви об'ємом десятки тисяч одиниць. Для оперативної оцінки параметрів розподілів і зменшення обчислювального навантаження доцільно використовувати дискретизовані ряди даних меншої розмірності, пропорційної часу реагування, для прийняття оперативних рішень. За таких умов постає питання про визначення потужності критеріїв оцінки експоненціальності динамічних рядів різного ступеня дискретизації, вибірок різних об'ємів.

Крім того, формування групи статистичних тестів, які доцільно використовувати при оцінці експоненціальності масивів даних ЩПН в комплексі «НБК–ОУ», проводилося з урахуванням їх інформаційної складової. Для створення автоматизованої системи оцінки та контролю експоненціальності динамічних рядів ЩПН, визначення моментів переходів до нестационарного стану, визначення параметрів критичності таких переходів пропонувані тести повинні мати одночасно і невелику обчислювальну ємність, і достатню точність, потужність та статистичну значущість.

Також група статистичних тестів, досліджуваних у роботі, включає можливість оцінки розподілів, які є версіями експоненціального – показникового, Вейбулла, гамма-розподілу. Оцінка потужності тестів проводилась на основі змодельованих динамічних рядів за різними законами розподілів з включенням експоненційної складової на різних часових проміжках.

Всі критерії оцінки експоненціального розподілу можна групувати за декількома ознаками, а саме: форма представлення даних для дослідження, вимоги до вибірки, просто-

та/складність розрахунків, потужність, формулювання гіпотез, метод представлення результату тощо.

Для розв'язання задачі визначення за реальним динамічним рядом спостережень моменту або часового інтервалу, коли закон розподілу набуває експоненціальної форми, доцільно визначити набір тестів, які враховують структуру досліджуваної вибірки, мають достатню статистичну потужність для визначених об'ємів вибірок, можуть бути порівняно легко розраховані та, за необхідності, програмно реалізовані. На нашу думку, для реалізації поставленої мети та задач дослідження доцільно використовувати перелік критеріїв, описаний нижче:

1) Критерій Шاپіро-Уїлка [9] можна використовувати для впорядкованої вибірки $x_1 \leq x_2 \leq \dots \leq x_n$, коли невідома початкова точка розподілу. При цьому розглядають щільність ймовірності з невідомим параметром μ :

$$f(x) = \frac{1}{v} \exp\left(-\frac{x-\mu}{v}\right). \quad (1)$$

Статистика Шاپіро-Уїлка для вибірки $n \leq 20$ має вигляд

$$W_E = \frac{n(\bar{x} - x_1)^2}{(n-1) \sum_{i=1}^n (\bar{x} - x_i)^2}, \quad (2)$$

де $n, x_i, \bar{x}$ – об'єм вибірки, i -те поточне та середнє значення рівнів динамічного ряду відповідно.

Для $n > 50$

$$W_E = \frac{(\bar{x} - x_1)^2}{\sum_{i=1}^n (\bar{x} - x_i)^2}. \quad (2.1)$$

Розраховані значення статистики W_E порівнюють з критичними $W_1(\alpha)$ та $W_2(\alpha)$ для заданого рівня значущості. Якщо виконується умова $W_1(\alpha) \leq W_E \leq W_2(\alpha)$, то гіпотеза про експоненціальність розподілу справедлива з достовірністю α . Модифікований критерій Шاپіро-Уїлка також може бути використаний для цензурованої (позбавленої від аномальних значень) впорядкованої вибірки.

2). Критерії типу Колмогорова-Смірнова включають критерій Андерсона-Дарлінга [10], критерій Смірнова-Крамера-фон Мізеса [11], критерій Уотсона [12]. Вони являють собою модифікації критерію Колмогорова-Смірнова для перевірки експоненціальності закону розподілу ймовірностей з невідомими параметрами:

$$F(x) = 1 - \exp\left(-\frac{x-\mu}{v}\right), \quad (3)$$

де μ та v – оцінки невідомих параметрів вибірки, які можуть бути знайдені із впорядкованої вибірки $x_1 \leq x_2 \leq \dots \leq x_n$ як

$$\begin{cases} \hat{v} = \frac{n(\bar{x} - x_1)}{n-1} \\ \hat{\mu} = x_1 - \frac{\hat{v}}{n} \end{cases} \quad (4)$$

Якщо провести заміну виду $\frac{x_i - \hat{\mu}}{\hat{v}} = w_i$, то розглядають нормований експоненціальний розподіл $z_i = 1 - \exp(-w_i)$. Тоді модифіковані критерії мають вигляд:

- Андерсона-Дарлінга:

$$A^2 = -\frac{1}{n} \sum_{i=1}^n (2i-1) [\ln z_i + \ln(1 - z_{n+1-i})] - n; \quad (5)$$

- Смірнова-Крамера-фон Мізеса:

$$W^2 = \sum_{i=1}^n \left(z_i - \frac{2i-1}{2n} \right)^2 + \frac{1}{12n}; \quad (6)$$

- Уотсона:

$$U^2 = W^2 - n(\bar{z} - 0,5)^2, \quad (7)$$

де $\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i$.

Розраховані значення статистик A^2, W^2 та U^2 порівнюють з критичними $A^2(\alpha), W^2(\alpha)$ та $U^2(\alpha)$ для заданого рівня значущості α . Якщо розраховані значення статистик менші критичних, то гіпотеза про експоненціальність розподілу справедлива з достовірністю α .

Окремо в цьому переліку слід звернути увагу на критерій Смірнова-Крамера-фон Мізеса для цензурованих даних [13]. При аналізі розглядається цензурована вибірка, в якій всі спостереження більші, ніж деяка величина x_p . Це означає, що проводиться оцінка експоненціальності k спостережень із вибірки об'ємом n , які задовольняють умові $0 \leq x_1 \leq x_2 \leq \dots \leq x_n$. Перевіряється гіпотеза про те, що k спостережень взяті із вибірки, яка має експоненційну функцію розподілу ймовірностей виду

$$F(x) = 1 - \exp\left(-\frac{x}{v}\right). \quad (8)$$

Оцінка параметра v визначається як

$$\hat{v} = \frac{1}{k} [\sum_{i=1}^n x_i + (n-k)x_n]. \quad (9)$$

Оцінка ступеня цензурування p :

$$\hat{p} = 1 - \exp\left(-\frac{x_n}{\hat{v}}\right). \quad (10)$$

При заміні $z_i = 1 - \exp\left(-\frac{x_i}{\hat{v}}\right)$ статистика критерію Смірнова-Крамера-фон Мізеса розраховується як

$$W_p^2 = \sum_{i=1}^k \left(z_i - \frac{2i-1}{2n}\right)^2 + \frac{1}{12k} \quad (11)$$

Якщо $W_p^2 \leq W_p^2(\alpha)$, гіпотеза експоненціальності приймається.

Приведені статистики оцінки експоненціальності ґрунтуються в основному на аналізі параметрів розподілу.

Наступна група статистик є достатньо простою в розрахунках і орієнтована на аналіз середніх величин, параметрів розсіювання, варіацій та зв'язків у досліджуваній вибірці.

3). Критерій Фроцині [14] може бути використаний до ранжованої, цензурованої чи неупорядкованої вибірки. Розрахунок статистики за критерієм простий і проводиться за формулою

$$B_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left| 1 - \exp\left(-\frac{x_i}{\bar{x}}\right) - \frac{i-0,5}{n} \right|, \quad (12)$$

де n , x_i , $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ відповідно об'єм вибірки, поточні значення та середнє арифметичне значення для вибірки.

Висновок про підтвердження чи відхилення гіпотез проводиться традиційно: якщо $B_n \leq B_n(\alpha)$ – гіпотеза експоненціальності приймається.

4). Критерій Фішера простий у розрахунках, може використовуватися до ранжованої, цензурованої чи неупорядкованої вибірки. Для вхідного ряду спостережень $x_1, x_2, \dots, x_n$ статистика критерію має вигляд

$$F = \frac{\sum_{i=1}^n x_i}{(n-1)x_i}. \quad (13)$$

Ця статистика при справедливості гіпотези про експоненційний розподіл у вибірці має F -розподіл Фішера з $f_1=2n-2$ та $f_2=2$ ступенями вільності. Якщо $F \leq F_{(\alpha, f_1, f_2)}$, то гіпотеза про експоненціальність приймається.

5). Критерій Бартлетта-Морана [15] також є достатньо простим у реалізації та не висуває окремих вимог до вибірки, яка аналізується. Для вхідного ряду $x_1, x_2, \dots, x_n$ статистика критерію має вигляд

$$B = \frac{12n^2}{7n+1} \left[\ln\left(\frac{1}{n} \sum_{i=1}^n x_i\right) + \frac{1}{n} \sum_{i=1}^n \ln x_i \right]. \quad (14)$$

При $n \geq 20$ розподіл статистики B добре апроксимується χ^2 -розподілом з $f = n - 1$ ступенями вільності. Якщо $B \leq \chi^2_{(\alpha, f)}$, то гіпотеза

за експоненціальності приймається з ймовірністю $1-\alpha$.

6). Критерій найбільшого інтервалу простий у реалізації і співпадає зі статистикою Кохрана для перевірки однорідності дисперсій. Статистика критерію визначається як

$$\eta_n = \frac{\max_{1 \leq i \leq n} (x_i - x_{i-1})}{\sum_{i=1}^n x_i}. \quad (15)$$

Якщо $\eta_n \leq \eta_n(\alpha)$, то гіпотеза про експоненціальність приймається.

7). Критерій Хартлі [16], або тест- $F_{\max}$ також використовується для перевірки однорідності дисперсій. Один із найбільш простих у реалізації і може використовувати нецензуровані дані. Статистика критерію визначається як

$$h(n) = \frac{\max_{1 \leq i \leq n} (x_i)}{\min_{1 \leq i \leq n} (x_i)}. \quad (16)$$

Якщо $h(n) \leq h(n)_\alpha$, то гіпотеза про експоненціальність приймається.

8). Становить інтерес кореляційний критерій експоненціальності [17], який має високу потужність і описує кореляційні зв'язки між рівнями ряду, що досліджується. При аналізі експоненційності розглядаємо функцію розподілу (3), оцінки параметрів вибірки μ та v розраховуються за формулами (4). Статистика критерію базується на коефіцієнті кореляції між нормованою змінною $\frac{x_i - \hat{\mu}}{\hat{v}} = w_i$ та математичним очікуванням i -ї порядкової статистики із експоненційного розподілу, який представлений вибіркою об'ємом n :

$$r_{(w,m)} = \frac{\sum_{i=1}^n (w_i - \bar{w})(m_i - \bar{m})}{\sqrt{\sum_{i=1}^n (w_i - \bar{w})^2 \sum_{i=1}^n (m_i - \bar{m})^2}}, \quad (17),$$

де $\bar{m} = \sum_{j=1}^i \frac{1}{n-j+1}$.

Якщо $n > 20$, то використовують $\tilde{m}_i = -\ln\left(1 - \frac{i}{n+1}\right)$. Тоді відповідні статистики мають вигляд

$$\begin{cases} K_{(w,m)} = n[1 - r_{(w,m)}^2] \\ K_{(w,\tilde{m})} = n[1 - r_{(w,\tilde{m})}^2] \end{cases} \quad (18)$$

Якщо $K_{(w,m)} \leq K_{(w,m)}(\alpha)$ чи $K_{(w,\tilde{m})} \leq K_{(w,\tilde{m})}(\alpha)$, гіпотеза експоненціальності приймається.

Серед критеріїв, які базуються на апроксимаціях неекспоненційних розподілів, слід відзначити критерій Климко-Антла-Редемакера-Рокетта [18] та критерій Лоулесса [19], які мають достатню потужність, прості у

розрахунках, байдужі до структури вибірки, легкі у програмній реалізації.

9). Критерій Климко-Антла-Редемакера-Рокетта базується на перевірці коефіцієнта форми $\beta=1$ у розподілі Вейбула з функцією виду

$$F(x) = 1 - \exp\left\{-\left(\frac{x-\mu}{\alpha}\right)^\beta\right\}. \quad (19)$$

Якщо $\beta=1$, то розподіл Вейбула переходить в експоненційний. Тому перевірка гіпотези $H_0: \beta=1$ еквівалентна перевірці гіпотези експоненціальності розподілу.

Для оцінки параметра β використовують апроксимацію $\tilde{c} \approx \nu^{-1,075}$, де $\nu = \frac{\sigma}{\bar{x}}$ – коефіцієнт варіації. В загальному вигляді для трьохпараметричного розподілу Вейбула

$$\nu = \frac{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}}{\frac{1}{n} \sum_{i=1}^n \left\{ x_i - \left[\min x_i - \frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \min x_i) \right] \right\}}. \quad (20)$$

Уточнене значення параметра c визначається як $c = \sqrt{n}(\tilde{c} - 1)$. Якщо $c < c_{(\alpha, n)}$, то гіпотеза $H_0: \beta=1$ (гіпотеза експоненціальності) – приймається.

10). Критерій Лоулесса базується на перевірці форми гамма-розподілу, статистика критерію визначається як

$$W = \frac{\tilde{x}}{\bar{x}}, \quad (21)$$

де $\tilde{x} = \sqrt[n]{\prod_{i=1}^n x_i}$ – середня геометрична для вибірки,

$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ – середня арифметична для вибірки

Якщо параметр форми гамма-розподілу дорівнює $k=1$, то розподіл переходить в експоненційний. Тому перевірка гіпотези $H_0: k=1$ проти альтернативних $H_1': k > 1$ та $H_1'': k < 1$ еквівалентна перевірці гіпотези про експоненціальний розподіл даних у вибірці, яка досліджується. Якщо виконується умова $W_1(\alpha) \leq W \leq W_{12}(\alpha)$, то гіпотеза про експоненціальність розподілу справедлива з достовірністю α .

Наведені вище статистичні критерії оцінки експоненціальності динамічних рядів обрані нами із великого масиву схожих критеріїв за ознаками простоти реалізації, достатньої достовірності результатів та можливості програмної реалізації. Із обраних критеріїв можна вичленити якісні групи, які мають деякі особливості їх використання, наведені в таблиці 1.

Таблиця 1 – Якісна характеристика критеріїв оцінки експоненціальності

№ групи	Перелік критеріїв	Загальна характеристика	Примітки
I	критерій Шапіро-Уїлка; критерій Андерсона-Дарлінга; критерій Смірнова-Крамера-фон Мізеса; критерій Уотсона	необхідність цензурування досліджуваної вибірки	
II	критерій Фроціні; критерій Фішера; критерій Бартлетта – Морана; критерій Хартлі	аналіз середніх; аналіз варіацій	використовуються для цензурованої та нецензурованої вибірок
III	EDF-критерій; критерій Климко-Антла-Редемакера-Рокетта; критерій Лоулесса	кореляційний аналіз; апроксимація неекспоненціальних розподілів	аналіз автокореляцій елементів вибірки; перевірка коефіцієнта форми в розподілі Вейбула; перевірка форми гамма-розподілу

Складено за: [3-19]

Методика та алгоритм експериментальних досліджень. У роботі множина аналізованих критеріїв досліджувалася методами імітаційного та статистичного моделювання. Оскільки робота є продовженням досліджень, опублікованих у [2], то для проведення імітаційного та статистичного моделювання та аналізу результатів обрані ті ж масиви спостережень динаміки ЦППН в ОУ за період 01.08.2019 р. – 31.08.2021 р.

Аналіз досліджуваних статистичних критеріїв оцінки експоненціальності включає імітаційне моделювання на дискретних вибірках кінцевої довжини, отриманих за допомогою експоненціальних та близьких до них законів розподілу, та перевірку результатів імітаційного моделювання на реальних вибірках. На етапі імітаційного моделювання рандомно були згенеровані вибірки кінцевої довжини для різних досліджуваних законів розподілу. Як і в роботі [2], для аналізу були обрані наступні об'єми досліджуваних вибірок: $n_1=730$ (дискретність – 1 день), $n_2=96$ (дискретність – 1 тиждень), $n_3=73$ (дискретність – 1 декада), $n_4=24$ (дискретність – 1 місяць). Вибір саме таких значень відповідає кумулятивним значенням ЦППН за день, тиждень, декаду та місяць.

За основу алгоритму досліджень взятий алгоритм, запропонований в [2], з модифікаціями відповідно до механізму аналізу:

1. За допомогою генератора випадкових чисел проводиться генерація масивів да-

них об'ємами $n_1=730$, $n_2=96$, $n_3=73$, $n_4=24$ для експоненційного розподілу (функція EXPON.DIST.), розподілу Вейбула (функція WEIBULL.DIST.), гамма-розподілу (функція GAMMA.DIST.) з відповідними налаштуваннями. При генерації масивів задані параметри середнього $\bar{x}_i$ та дисперсії σ_i^2 були максимально наближені до реальних значень середнього та дисперсії ЩПН на досліджуваному об'єкті за період 01.08.2019 р. – 31.08.2021 р.

2. Для критеріїв I групи (таблиця 1) проводиться цензурування змодельованої вибірки за зростанням рівнів ряду за зростанням кожної згенерованої вибірки; для критеріїв груп I-III проводиться визначення оцінок невідомих параметрів вибірки $\bar{x}_i$, σ_i^2 , μ та ν .

3. Розрахунок статистик критеріїв та формулювання висновків про відповідність/невідповідність гіпотезі, яка перевіряється.

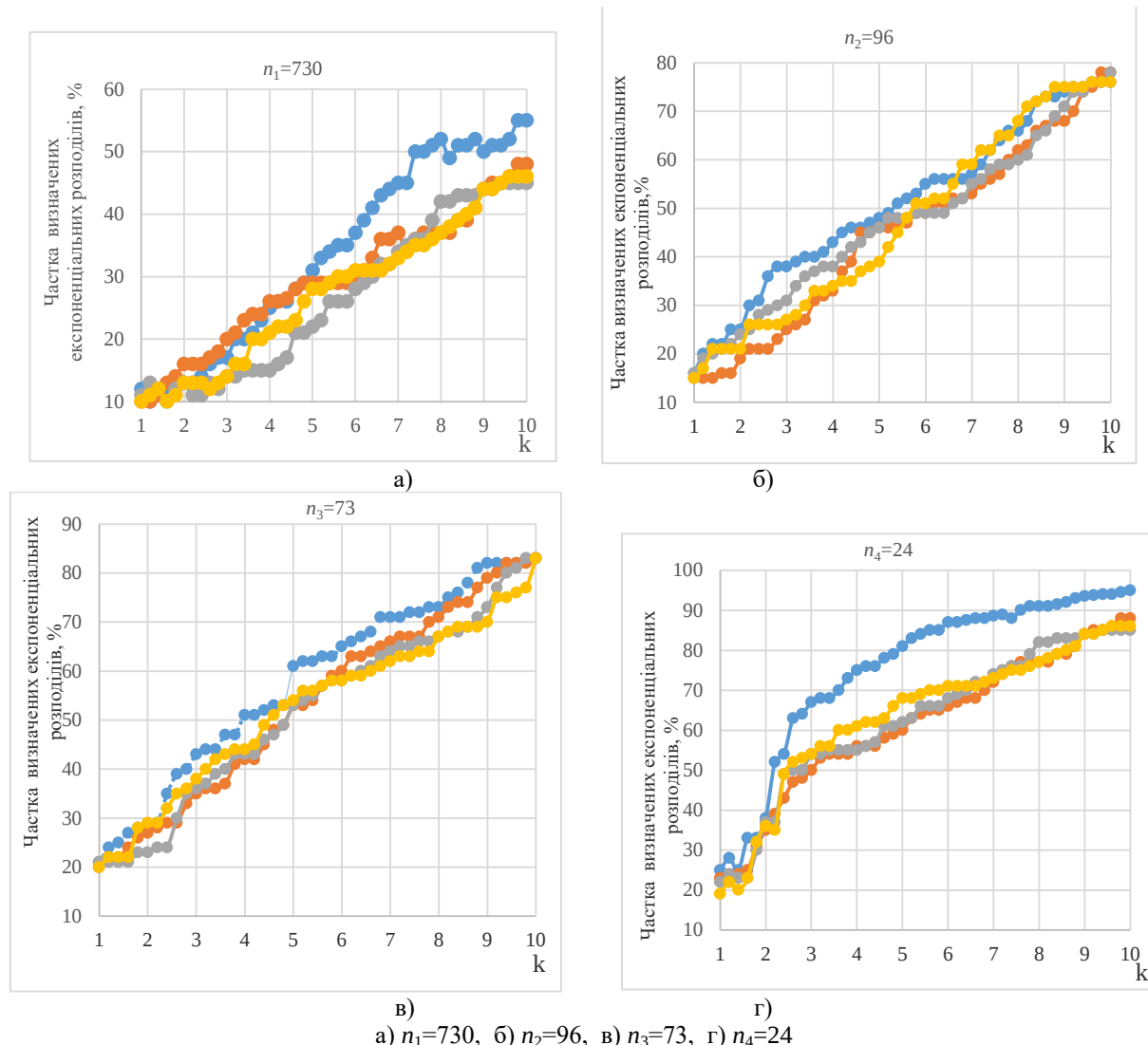
4. Визначення потужностей критеріїв як доли правильно ідентифікованих розподілів у масиві експериментальних даних.

5. Аналіз потужності критеріїв залежно від величини вхідної вибірки.

Слід зазначити, що для визначення потужностей критеріїв для кожного окремого об'єму вибірки n_i за заданим законом розподілу проводилося генерування $N=10$ вибірок за умовою $\bar{x}_i \pm k\sigma_i$, де σ_i – середнє квадратичне відхилення i -ї вибірки, $k = [1,0 \div 10,0]$.

Потужність критерію визначається як доля правильно ідентифікованих законів розподілу для рівня значущості $\alpha=0,05$.

Результати імітаційного моделювання та моніторингу потужностей критеріїв. Результати моніторингу критеріїв на основі згенерованих масивів даних для першої групи критеріїв зображено на рисунку 1.



а) $n_1=730$, б) $n_2=96$, в) $n_3=73$, г) $n_4=24$

● – критерій Шапіро-Уїлка; ● – критерій Андерсона-Дарлінга;
● – критерій Смірнова-Крамера-фон Мізеса; ● – критерій Уотсона

Рисунок 1 – Оцінка потужностей першої групи критеріїв

Результати моніторингу критеріїв на основі згенерованих масивів даних

для другої групи критеріїв зображено на рисунку 2.

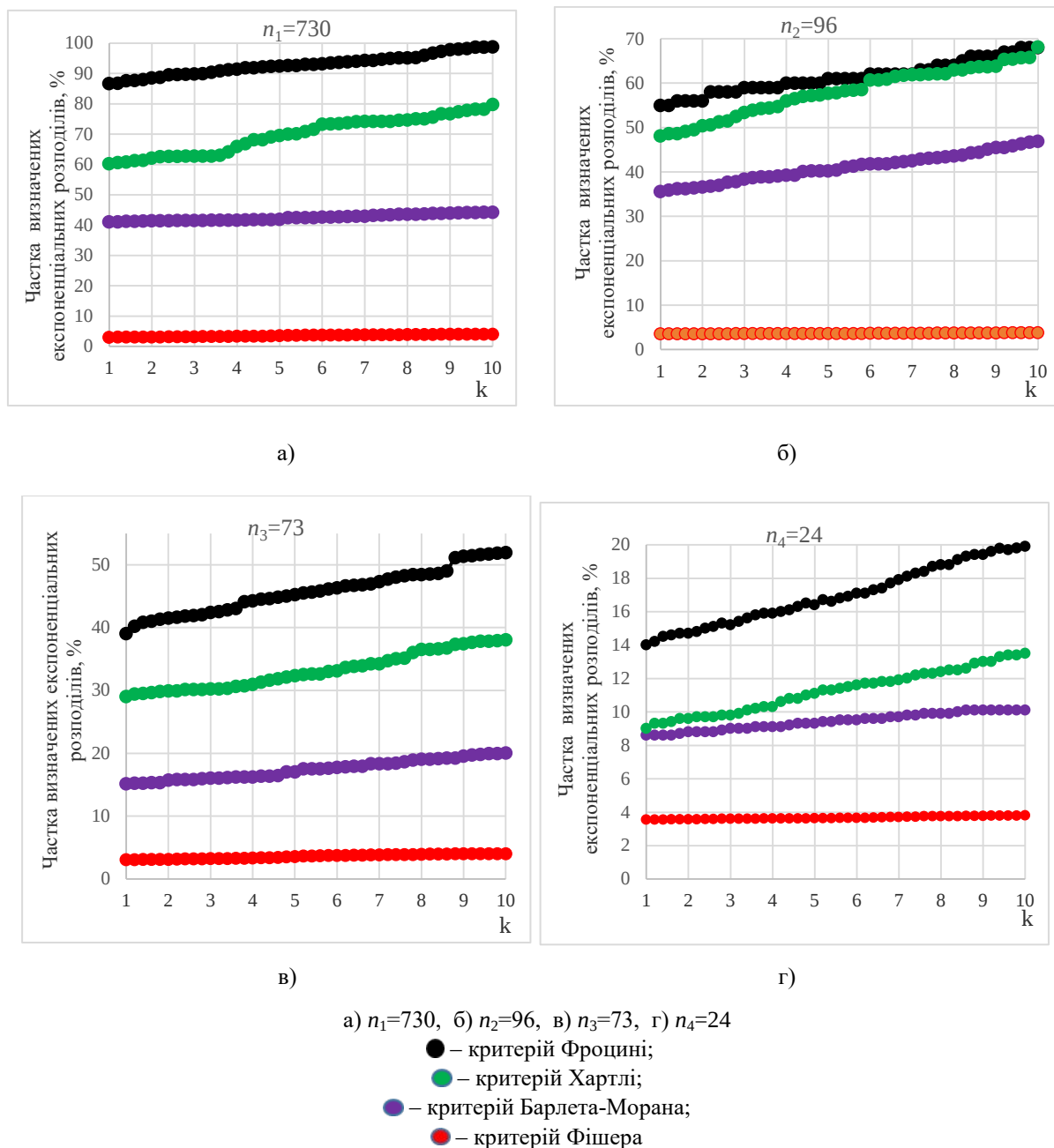


Рисунок 2 – Оцінка потужностей другої групи критеріїв

Результати імітаційного моделювання та моніторингу потужностей статистичних

критеріїв тестування експоненціальності III групи зображено на рисунку 3.

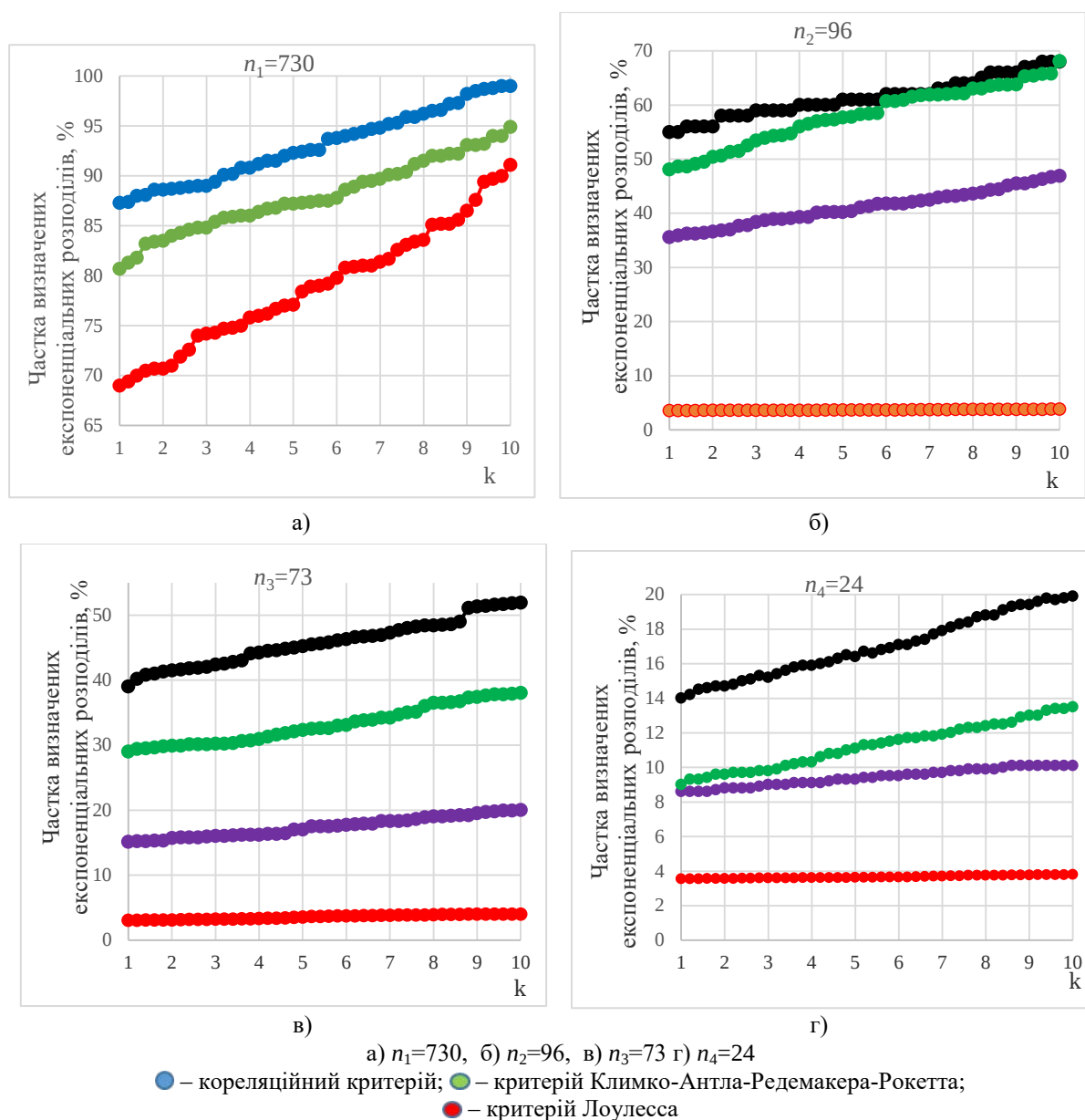


Рисунок 3 – Оцінка потужностей третьої групи критеріїв

Обговорення результатів. Проведений моніторинг потужностей критеріїв дає змогу стверджувати, що перша група критеріїв демонструє гарну потужність на невеликих вибірках ($n_4=24$), яка зменшується зі зростанням об'ємів вибірок ($n_1=730$, $n_2=96$, $n_3=73$). Слід зазначити також, що зі зростанням рівня дисперсії вибірок $\bar{x}_i \pm k\sigma_i$, які досліджувалися, така тенденція щодо величини вибірок зберігається на всьому діапазоні досліджень. Також із проведених досліджень видно, що критерії типу Колмогорова-Смірнова поступаються за потужністю критерію Шапіро-Уїлка на всіх діапазонах значень x_i .

Друга група критеріїв демонструє високу потужність на великих вибірках ($n_1=730$,

$n_2=96$), за винятком критерію Фішера. Критерії Фроцини та Хартлі мають значно більшу потужність порівняно з критеріями Фішера та Барлета-Морана. Потужність критеріїв зменшується зі зменшенням об'ємів вибірок. Зростання рівня дисперсії вибірок $\bar{x}_i \pm k\sigma_i$ практично не впливає на рівень потужності критеріїв, що свідчить про їх сталість та значущість на всьому діапазоні значень x_i .

Третя група критеріїв ефективна для вибірок великих об'ємів ($n_1=730$), проте потужність критеріїв значно (приблизно в 2-4 рази) зменшується при зменшенні об'ємів вибірок ($n_2=96$, $n_3=73$, $n_4=24$). Збільшення об'ємів вибірок має вплив на потужності кри-

теріїв. На вибірках великих об'ємів цей вплив становить 1,35-2,3% на $k\sigma_i$.

Слід зазначити, що потужності критеріїв визначалися для рівня значущості $\alpha=0,05$.

Висновки. Другий етап дослідження динаміки ЩПН в комплексі «НБК-ОУ» включає задачі визначення моментів / інтервалів переходу трендів до експоненціальних законів розподілу. Проведене дослідження дає змогу констатувати, що при розв'язку задачі трендового аналізу масивів даних ЩПН в системі контролю ядерної безпеки ПВМ на другому етапі необхідно визначити набір оптимальних критеріїв оцінки експоненціальності та меж їх використання. В результаті імітаційного моделювання встановлено, що максимальна ефективність статистичного оцінювання параметрів експоненціальності досягається при використанні для тестування критеріїв з максимальною потужністю, яка, в свою чергу, залежить від величини досліджуваної вибірки.

Так, для масивів спостережень з рівнем дискретизації 1 день (об'єм вибірки $n_1=730$) максимально ефективним буде використання критерію Фроціні (II група) та кореляційного критерію експоненціальності і критерію Климо-Антла-Редемакера-Рокетта (III група), потужність яких лежить в межах 87-99%.

Для масивів спостережень з дискретизацією 1 тиждень ($n_2=96$) доцільно використовувати критерії I групи, максимальна потужність яких (50-75%) досягається на динамічних рядах зі значним рівнем дисперсії $\bar{x}_i \pm k\sigma_i$, при $k>5$.

Вибірки з рівнями дискретизації 1 декада ($n_3=73$) оцінюються на експоненціальність максимально ефективно за допомогою критеріїв I групи з потужностями 50-85% для $\bar{x}_i \pm k\sigma_i$, при $k>5$; для вибірок з дискретизацією 1 місяць ($n_4=24$) – критерій Шапіро-Уїлка (потужність критерію становить 63-95% на діапазонах $k>3\sigma$).

Наукова новизна отриманих результатів полягає в уточненні переліку критеріїв тестування на експоненціальність для дискретизованих динамічних рядів ЩПН, конкретизації та класифікації меж їх використання, використанні визначених оптимальних наборів критеріїв моніторингу на нові класи об'єктів комплексу «НБК-ОУ».

Практична значущість полягає в тому, що проведені дослідження та отримані результати є основою для проектування і створення підсистеми тестування та моніторингу експоненціальності в автоматизованій

системі контролю ЩПН та автоматизованій системі підтримки прийняття рішень по забезпеченню та підвищенню рівня ядерної безпеки комплексу «НБК ОУ».

Перспективою подальших досліджень є використання отриманих результатів разом з визначеними в [2] наборами тестів оцінки аномальності в дискретизованих масивах даних ЩПН для побудови алгоритмів функціонування автоматизованих систем контролю комплексу «НБК ОУ».

Список використаних джерел

- [1] І. Г. Шараєвський, Т. С. Власенко, Л. Б. Зімін, А. В. Носовський, Н. М. Фіалко, та Г. І. Шараєвський, "Перспективні напрями підвищення експлуатаційної надійності та забезпечення оперативного управління ресурсом головного обладнання", *Ядерна енергетика та довкілля*, т. 3, № 22, с. 3-13, 2021.
- [2] І. С. Скітер, та М. В. Савельєв, "Моделювання аномалій щільності потоку нейтронів в автоматизованій системі контролю ядерної безпеки об'єкта "Укриття", *Математичні машини і системи*, № 4, с. 70-77, 2021.
- [3] А. І. Кобзар, *Прикладна математична статистика. Для інженерів і наукових робітників*. ФІЗМАТЛІТ, 2006.
- [4] B. V. Frozini, *On the Distribution and Power of a Goodness-of-fit Statistic with Parametric and Nonparametric Applications*, "Goodness-of-fit". North-Holland, 1987.
- [5] A. Vahideh, and H. Arezou, "Test based on progressively type-II censored data via extension of cumulative Tsallis divergence", *Revista Colombiana de Estadística - Applied Statistics*, vol. 44, pp. 97-312, 2021.
- [6] M. Mansour, "Non-parametric statistical test for testing exponentiality with applications in medical research", *Statistical Methods in Medical Research*, vol. 29, no. 2, pp. 413-420, 2020.
- [7] A. Adhikari, and J. Schaffer, "Modified Lilliefors test", *Journal of Modern Applied Statistical Methods*, vol. 14, no. 1, pp. 53-69, 2015.
- [8] E. O. Ossai, M. S. Madukaife, and A. V. Oladugba, "A review of tests for exponentiality with Monte Carlo comparisons", *Journal of Applied Statistics*, no. 2, pp. 345-420, 2020.
- [9] S. S. Shapiro, and M. B. Wilk, "An analysis of variance test for the exponential distribu-

- tion (complete samples)", *Technometrics*, vol. 14, no. 1, pp. 355-370, 1972.
- [10] T. W. Anderson, and D. A. Darling, "Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes", *Annals of Mathematical Statistics*, vol. 23, pp. 193-212, 1952.
- [11] T. W. Anderson, "On the distribution of the two-sample Cramer-von Mises criterion", *Annals of Mathematical Statistics. Institute of Mathematical Statistics*, vol. 33, no. 3, pp. 1148-1159, 1962.
- [12] J. D. Spurrier, "An overview of tests for exponentiality", *Communications in Statistics - Theory and Methods*, vol. 13, pp. 56-69, 1984.
- [13] D. A. Darling, "The Kolmogorov-Smirnov, Cramer-von Mises tests", *The Annals of Mathematical Statistics*, vol. 28, no. 4, pp. 823-838, 1957.
- [14] R. Mezbahur, and W. Han, "Tests for exponentiality: A comparative study", *American Journal of Applied Mathematics and Statistics*, vol. 5, no. 4, pp. 125-135, 2017.
- [15] M. S. Barlett, "Properties of sufficiency and statistical tests", *Proceedings of the Royal Statistical Society*, vol. 160, pp. 268-282, 1937.
- [16] H. O. Hartley, "The maximum F-ratio as a short cut test for homogeneity of variance", *Biometrika*, vol. 37, pp. 308-312, 1950.
- [17] M. A. Stephens, *EDF-Tests of Fit for the Logistic Distribution: Technical report*. Department of Statistics, Stanford University (Prepared under Grant DAAG29-77-G-0031 for the U.S. Army Research Office), no. 36, 1979.
- [18] L. A. Klimko, C. E. Antle, A. W. Rademaker, and H. E. Rockette, "Upper bounds for the power of invariant tests for the exponential distribution with Weibull alternatives", *Technometrics*, vol. 17, no. 3, pp. 357-360, 1975.
- [19] J. F. Lawless, *Statistical Models and Methods for Lifetime Data* (2-nd ed.). Hoboken, N.-J., USA: John Wiley & Sons, Inc., 2003.
- [20] I. S. Skiter, and M. V. Saveliev, "Simulation of neutron flux density anomalies in the automated nuclear safety control system of the Shelter's object", *Matematychni mashyny i systemy*, no. 4, pp. 70-77, 2021 [in Ukrainian].
- [21] A. I. Kobzar, *Applied Mathematical Statistics. For engineers and researchers*. FIZMATLIT [in Ukrainian].
- [22] B. V. Frozini, *On the Distribution and Power of a Goodness-of-fit Statistic with Parametric and Nonparametric Applications, "Goodness-of-fit"*. North-Holland, 1987.
- [23] A. Vahideh, and H. Arezou, "Test based on progressively type-II censored data via extension of cumulative Tsallis divergence", *Revista Colombiana de Estadística - Applied Statistics*, vol. 44, pp. 97-312, 2021.
- [24] M. Mansour, "Non-parametric statistical test for testing exponentiality with applications in medical research", *Statistical Methods in Medical Research*, vol. 29, no. 2, pp. 413-420, 2020.
- [25] A. Adhikari, and J. Schaffer, "Modified Lilliefors test", *Journal of Modern Applied Statistical Methods*, vol. 14, no. 1, pp. 53-69, 2015.
- [26] E. O. Ossai, M. S. Madukaife, and A. V. Oladugba, "A review of tests for exponentiality with Monte Carlo comparisons", *Journal of Applied Statistics*, no. 2, pp. 345-420, 2020.
- [27] S. S. Shapiro, and M. B. Wilk, "An analysis of variance test for the exponential distribution (complete samples)", *Technometrics*, vol. 14, no. 1, pp. 355-370, 1972.
- [28] T. W. Anderson, and D. A. Darling, "Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes", *Annals of Mathematical Statistics*, vol. 23, pp. 193-212, 1952.
- [29] T. W. Anderson, "On the distribution of the two-sample Cramer-von Mises criterion", *Annals of Mathematical Statistics. Institute of Mathematical Statistics*, vol. 33, no. 3, pp. 1148-1159, 1962.
- [30] J. D. Spurrier, "An overview of tests for exponentiality", *Communications in Statistics - Theory and Methods*, vol. 13, pp. 56-69, 1984.
- [31] D. A. Darling, "The Kolmogorov-Smirnov, Cramer-von Mises tests", *The Annals of Mathematical Statistics*, vol. 28, no. 4, pp. 823-838, 1957.

References

- [1] I. G. Sharaevskiy, T. S. Vlasenko, L. B. Zimin, A. V. Nosovskiy, N. M. Fialko, and G. I. Sharaevskiy, "Perspective directions of increase of operational reliability and maintenance of operative management by a resource of the main equipment", *Yaderna*

- [14] R. Mezbahur, and W. Han, "Tests for exponentiality: A comparative study, *American Journal of Applied Mathematics and Statistics*, vol. 5, no. 4, pp. 125-135, 2017.
- [15] M. S. Barlett, "Properties of sufficiency and statistical tests", *Proceedings of the Royal Statistical Society*, vol. 160, pp. 268-282, 1937.
- [16] H. O. Hartley, "The maximum F-ratio as a short cut test for homogeneity of variance", *Biometrika*, vol. 37, pp. 308-312, 1950.
- [17] M. A. Stephens, *EDF-Tests of Fit for the Logistic Distribution: Technical report*. Department of Statistics, Stanford University (Prepared under Grant DAAG29-77-G-0031 for the U.S. Army Research Office), no. 36, 1979.
- [18] L. A. Klimko, C. E. Antle, A. W. Rademaker, and H. E. Rockette, "Upper bounds for the power of invariant tests for the exponential distribution with Weibull alternatives", *Technometrics*, vol. 17, no. 3, pp. 357-360, 1975.
- [19] J. F. Lawless, *Statistical Models and Methods for Lifetime Data* (2-nd ed.). Hoboken, N.-J., USA: John Wiley & Sons, Inc., 2003.

I. S. Skiter¹, Ph. D., Associate Professor,
e-mail: skiteris@ukr.net

M. V. Saveliev^{1,2}, Ph. D.,
e-mail: m.saveliev@isnpp.kiev.ua

S. V. Kupriyanchuk¹,

D. O. Khomenko^{1,2}

e-mail: khomenko.amid@gmail.com

¹ Institute for Safety Problems of Nuclear Power Plants, NAS of Ukraine
Kirova st., 36-a, Chornobyl, 07270, Ukraine

² Institute of Mathematical Machines and Systems Problems of the Ukraine NAS
acad. Glushkova ave., 42, Kyiv, 03187, Ukraine

ALGORITHM FOR MONITORING THE POWER OF CRITERIA FOR TESTING THE EXPONENTIVITY OF NEUTRON FLUX DENSITY IN THE SHELTER OBJECT

Nuclear safety monitoring of the "New Safe Confinement – Shelter Object" system requires constant monitoring of the dynamics of neutron flux density from clusters of nuclear fuel-containing materials at sub-reactor premises at the Shelter object. The paper proposes an algorithm for identifying transitions of neutron flux time series trends from lineal into exponential dependence. The monitoring of the powers of statistical criteria that can be used in determining the moments of transition of dynamic series to the law of exponential distribution has been carried out. The criteria are classified according to the "power - sample size" parameters. A set of criteria, which is most expedient to use when checking the laws of exponential distribution, is offered. The power of statistical criteria is estimated at different sample sizes, without taking into account competing hypotheses. Powers of three groups of exponentiality testing criteria for different data sets have been established. Motivated use of criteria for different sample sizes is recommended. The obtained results are the basis for the construction of algorithms for automated monitoring systems for assessing the state of fuel-containing materials in the complex "New Safe Confinement – Shelter Object". The scientific novelty of the obtained results consists in clarifying the list of exponentiality testing criteria for discretized time series, specifying the limits of their use, determining optimal sets of monitoring criteria for new classes of objects. The practical significance lies in the use of the results as a basis for designing a subsystem for exponentiality testing and monitoring in an automated control system and decision support system on the research object. The direction of further research is the use of the obtained results to construct algorithms for automated monitoring and decision support systems to assess the state of the New Safe Confinement – Shelter Object system.

Keywords: "New Safe Confinement – Shelter Object", nuclear-hazardous clusters, simulation modeling, time series analysis, operational diagnostics, hypothesis testing, statistical tests, decision support.

[0000-0002-5387-7717] **І. Г. Зінченко**, здобувач осв.-наук. ступеня доктора філософії,
e-mail: igor.zinchenko557@gmail.com

[0000-0001-5571-2281] **О. В. Лавданська**, канд. техн. наук, доцент

Черкаський державний технологічний університет
б-р Шевченка, 460, м. Черкаси, 18006, Україна

СУЧАСНІ ТЕХНОЛОГІЇ ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ ДІДЖИТАЛІЗАЦІЇ

У статті досліджуються сучасні технології оцінювання ефективності діджиталізації. Проаналізовано процес діджиталізації. Детально розглянуто такі переваги цифрової трансформації, як оптимізація бізнес-процесів, підвищення прибутковості компанії та клієнтоцентризм. Визначено особливості проєкту із впровадження діджиталізації. Наведено опис задач і проблем, актуальних для процесу впровадження діджиталізації. Зазначені проблеми та виклики мають місце у будь-якій предметній області. Виконано аналіз поточного стану технологій та методів оцінки ефективності діджиталізації. Визначено можливі варіанти реалізації механізму прийняття рішень при оцінці ефективності діджиталізації. Описано особливості, принципи застосування та базові елементи інтелектуальних інформаційних систем оцінювання. Наведено приклад структури інтелектуальної системи, що вирішує задачі оцінювання впроваджених заходів діджиталізації. До елементів інформаційної системи оцінки ефективності діджиталізації можуть належати: зацікавлені особи, апаратне забезпечення, експерти, комплекс прийняття рішень, база знань, інтерфейс користувача. Визначено можливі напрямки подальших досліджень.

Ключові слова: цифровізація, оцінювання, ефективність, система, автоматизація, прийняття рішень.

Вступ. У сучасному світі кількість цифрової інформації зростає все швидше з кожним днем, що впливає і на цифровізацію, тому все більше постає питання оцінки ефективності впроваджених заходів. Інтелектуальні комп'ютерні системи можуть бути корисними у цьому процесі, оскільки при коректному налаштуванні вони можуть точно та швидко дати результат, а також вони зводять до мінімуму вплив людського фактора на процес оцінки.

Аналіз останніх досліджень та публікацій. Ефективність впровадження діджиталізації в сферах людської діяльності досить широко досліджується в усьому світі. Наприклад, Богавач Міланка (Bogavac Milanka) демонструє новий метод оцінки діджиталізації малого та середнього бізнесу – індекс IDSME, який передбачає самооцінювання суб'єкта для виявлення слабких сторін та створення плану розвитку [1]. Індекс IDSME є розвитком Digital Economy and Society Index (DESI), що застосовується в ЄС для оцінки впроваджених заходів цифровізації за допомогою кількох індикаторів [2]:

- тип підключення;
- цифрові навички;
- використання Інтернету;

- впровадження цифрових технологій;
- публічні цифрові сервіси.

Команда дослідників із Азербайджану [3] дослідила та оцінила вагомість критеріїв ефективності діджиталізації на прикладі сфери сільського господарства. Використовуючи результати опитування серед 1513 топ-менеджерів сільськогосподарських компаній, вони дійшли висновку, що ключовими обмежувачами факторами цифровізації є вартість впровадження та використання цифрових технологій, законодавчо не регульовані процеси цифровізації галузі, а також рівень кібербезпеки. Надійність результатів опитування перевірялась за допомогою альфи Кронбаха, а вагові коефіцієнти кожного фактора оцінювались за допомогою інтегральних моделей та показника інформаційної ентропії. Використовуючи ті ж методи, дослідники також вивели окремий показник загального рівня діджиталізації, за рахунок якого вдалося побудувати шкалу розподілу господарств з трьох основних рівнів та двох перехідних. Результати оцінки показали, що приріст прибутковості від впроваджених заходів при низькому рівні діджиталізації дорівнює 1,08 % при підвищенні загального показника на 1 %, а для господарств

із середнім рівнем розвитку приріст становить 1,59 %. У дослідженні відсутні компанії високого рівня діджиталізації, проте можна припустити, що для них результати можуть бути ще вищими.

Дослідження [4] пропонує метод оцінки діджиталізації індексом DFSI в контексті розвитку фінансового сектору європейських країн. Якщо узагальнити його ключові складові, то індекс включає такі компоненти: індикатор цифровізації, розвиток відповідного ринку послуг, ступінь впровадження діджиталізації при наданні послуг. Для підрахунку DFSI використовуються загальнодоступні статистичні індекси, кожен з яких висвітлює конкретний напрямок використання цифрових технологій у галузі. Для безпосереднього розрахунку індексу використовується метод зважених сум на основі вагових коефіцієнтів для кожного з підіндексів. Для їх обчислення використовувалась формула Фішборна, оскільки для нього достатньо лише знати пріоритет підіндексів між собою. Як і в попередньому [3] дослідженні, було побудовано шкалу із чотирьох інтервалів для якісного відображення даних.

В іншій публікації [5] пропонується використання розробленого індексу DPI для оцінки рівня цифровізації морських портів. Дослідники розробили систему з 22 критеріїв, яка була поділена на чотири групи, кожна з яких має свій ваговий коефіцієнт. Далі представники 30 європейських портів оцінили зазначені фактори за шестибальною шкалою Лайкерта протягом різних проміжків часу. Через наявність інформаційного шуму до результатів опитування було застосовано фільтр Калмана, який мінімізує цей фактор. Як і у вищезазначених дослідженнях, у цьому також було створено інтервальну шкалу розподілу результатів розрахунку індексу для якісного відображення даних.

Дослідник пропонує [6] використовувати граф-теоретичну модель для оцінки ефективності використання різних інформаційних технологій у рамках діджиталізації. Вершинами в подібному графі є різні стани системи та оточення в цілому, а ребрами графу є операції, що здійснюються в цій обраній системі. Для побудови такої моделі необхідно спершу створити концептуальні моделі використання інформаційних технологій, на основі яких будуються алгебраїчні графи. Далі графи помічаються змінними та параметрами, а також позначаються функціональні залежності. До

позначок функціональних залежностей може додаватися програмний код. Якщо об'єднати послідовність таких функціональних змін станів з кодом, то в результаті отримаємо програмну модель системи. Таким чином імітується робота реальної системи з використанням інформаційних технологій.

Огляд наявних джерел показав, що тема оцінки діджиталізації та її ефективності є досить поширеною в різних сферах людської діяльності. Широко використовуються методи багатокритеріальної оптимізації для пошуку значень розроблених індексів, як-то метод зважених сум. Також можна зробити висновок про те, що досить часто використовуються експертні опитування як інструмент для прийняття рішень.

Проведення власних досліджень у цій області може допомогти формалізувати і систематизувати загальні критерії ефективності цифровізації та її оцінку, розробити структурні й математичні моделі роботи подібних інтелектуальних систем, що працювали б незалежно від предметної області, а також перевірити їх на практиці. В рамках цих досліджень доцільно використати вищезазначені методи, оскільки вони показали свою ефективність.

Мета та задачі дослідження. Метою дослідження є аналіз процесу діджиталізації та аналіз технологій, які використовуються в процесах оцінювання ефективності діджиталізації, для визначення впливу діджиталізації на результати діяльності компаній шляхом побудови інтелектуальної системи оцінювання ефективності діджиталізації.

Задачі дослідження передбачають:

- формулювання особливостей, проблем та задач впровадження діджиталізації;
- аналіз підходів до оцінювання ефективності впровадження діджиталізації в різних сферах людської діяльності;
- визначення структурно-логічних елементів побудови інтелектуальних систем оцінювання ефективності впровадження діджиталізації.

Зміст впровадження діджиталізації. В більшості випадків необхідність перебудови бізнес-моделі компанії зумовлюється низькою ефективністю компанії. Про це свідчить зменшення успішності та відсутність необхідних результатів у компанії.

Невід'ємною частиною успішних компаній є безперервне переосмислення способу розвитку компанії, спрямованого на простий,

економічний і надійний варіант роботи, оновлення та дотримання нових напрямків на ринку технічних розробок.

Впровадження діджиталізації полягає в оптимізації певних аспектів людської діяльності за допомогою інформаційно-комунікаційних технологій. Серед таких аспектів є [3]:

- документообіг;
- управлінські процеси;
- прийняття рішень;
- конкурентоспроможність;
- час працівників, їх саморозвиток.

Вищезазначені аспекти будуть актуальними для більшості галузей, проте для кожної окремої з них можуть бути додаткові процеси, що можуть бути покращені за рахунок цифровізації.

Процес діджиталізації стимулює зміни в бізнес-середовищі. Основою для таких змін є впровадження нових технологій, таких як: аналіз великих даних (Big Data), Інтернет речей, хмарні технології, штучний інтелект, розпізнавання голосу, 3D-друк, віртуальна та доповнена реальності тощо. Таким чином, відбувається перехід від класичної «продуктової» організації до технологічної, застосування нових управлінських моделей, що ґрунтуються на розробці digital-стратегії. Завдяки таким змінам підвищується ефективність бізнес-процесів [7].

Діджиталізація бізнесу – це еволюційний процес впровадження знань і творчості в бізнесі. Оптимізація бізнесу шляхом впровадження IT-рішень та роботизації включає нові зовнішні фактори та можливості, які потребують перегляду стратегічних пріоритетів. Необхідність у цифровій трансформації зумовлюється прагненням до постійної адаптації бізнес-середовища до потреб споживачів, а з погляду компаній – покращення власної конкурентоспроможності на ринку. У цифровому світі досягають успіху ті компанії, які прагнуть постійних змін та здатні запропонувати цінний для споживача продукт. Розвиток цифрових технологій допомагає удосконалювати діяльність компанії у різних напрямках: маркетинг, обслуговування клієнтів, логістика, внутрішні процеси тощо, впливаючи на активність зростання доходів та підвищення конкретних позицій компанії. Адже цифрова трансформація сприяє задоволенню потреби власника бізнесу захистити своїх співробітників у нетипових умовах роботи, дозволяє організувати віддалено спільну роботу над проєктами,

проводити навчання, запобігати збоєм та зупинкам у роботі шляхом автоматизації процесів та роботизації.

За підрахунками «Українського інституту майбутнього», реалізація цільового сценарію переходу української економіки до цифрової за 3–5 років дозволить [8]:

- створювати щонайменше від 11 % (у 2021E) до 95 % (2030E) додаткового ВВП на рік;
- за 10 років додатково створити до \$1260 млрд ВВП;
- за 10 років збільшити надходження в бюджет на \$240 млрд;
- створити 700 тис. нових робочих місць (не враховуючи експортну IT-індустрію).

Розглянемо основні переваги цифрової трансформації:

1. Оптимізація бізнес-процесів сприяє ефективному просуванню бізнесу, зокрема, дозволяє:

- заощадити час і підвищити продуктивність праці;
- автоматизувати виробництва, автоматизувати рутинні процеси й інші внутрішні процеси компанії;
- оптимізувати та покращити внутрішні і зовнішні комунікації;
- розширити цільову аудиторію, канали та географію продажу;
- корегувати діяльність компанії завдяки аналітичним даним.

2. Підвищення прибутковості компанії шляхом впровадження нових технологічних змін та їх діджиталізація позитивно позначаються на:

- збільшенні продажів;
- збільшенні фінансової вигоди шляхом скорочення нерентабельних вкладень;
- скороченні бюджету на реалізацію маркетингових стратегій без шкоди для ефективності просування.

3. Клієнтоцентризм. Клієнти є джерелом постійного покращення та розвитку для компанії. Оперативне реагування на потреби клієнта здатне забезпечити:

- можливості крос-продажів / upsell-продажів;
- вихід на новий рівень обслуговування клієнтів і заохочення їх до частої продуктивної співпраці з компанією;
- зменшення частки втрачених клієнтів та мінімізацію помилок під час обслуговування клієнтів;

- підвищення конкурентних можливостей компанії, підвищення лояльності клієнтів до компанії та покращення клієнтського досвіду завдяки персоналізованому підходу до роботи з клієнтами;
- скорочення шляху клієнта від звернення у компанію до отримання готового продукту (надання послуги) шляхом надання клієнту нових можливостей: відстеження всього ланцюжка співпраці та використання роботизованих інструментів для консультацій 24/7.

Особливості впровадження діджиталізації. Діджиталізація в кожному окремому випадку є інноваційним проєктом. Можна назвати наступні етапи такого проєкту:

- аналіз поточних ресурсів та процесів, під час якого визначаються пріоритетні напрямки для змін;
- розробка плану переходу до використання нових технологій разом з кінцевими цілями;
- безпосередньо реалізація запланованих заходів;
- оцінка та систематизація отриманих результатів.

Причому, такий проєкт у рамках одного підприємства може бути реалізований довільну кількість разів, що буде залежати від наявності аспектів, які ще можна оптимізувати, або ж у випадку модернізації вже наявних цифрових засобів.

Задачі та проблеми впровадження діджиталізації. Відомо, що кожна предметна область має свої особливості функціонування, проте існують загальні задачі або цілі цифровізації, які можна застосувати до будь-якої з них. Виходячи з самого поняття діджиталізації, задачі полягають у:

- розробці та реалізації плану переходу до нової парадигми;
- міграції потоків інформації в цифровий вигляд;
- зменшенні або ліквідації людського фактора в типових задачах;
- автоматизації процесів, які можна автоматизувати.

Тобто загалом процес цифровізації має на меті всебічну оптимізацію бізнес-процесів та адаптацію суб'єкта до нових реалій епохи четвертої промислової революції.

Якщо формулювати проблеми, виходячи з задач, то можна навести такі:

- недовіра бізнесу до інформаційних технологій;

- визначення некоректних або неефективних критеріїв оцінки впроваджених заходів;
- постановка занадто абстрактних цілей у підході до процесу діджиталізації;
- неготовність певних груп клієнтів до впровадження нових технологій;
- необхідність завчасного продумування та побудови мапи взаємодії з клієнтом, ефективною для обох зацікавлених сторін;
- ретроспектива невдалих спроб трансформації;
- нехтування або недостатня увага до кібербезпеки.

Останній пункт є особливо важливим, оскільки неправомірний доступ до інформаційних ресурсів може мати ризик як для самої інформації, з погляду її конфіденційності, так і для інформаційних систем, що залучені до бізнес-процесів компанії.

Компанія Panorama Consulting Group наводить такі виклики [9]:

- брак кваліфікованих ІТ-спеціалістів;
- брак управління організаційними змінами;
- зміни потреб споживачів;
- брак визначеної стратегії;
- проблеми та обмеження бюджету;
- неефективне керування даними;
- неефективні бізнес-процеси.

У той же час компанія gotodigital [10] називає брак методології причиною 85 % невдач у процесі діджиталізації.

Аналіз технологій, які застосовуються в процесах оцінки діджиталізації. Автор припускає, що для подібних потреб можна застосувати інтелектуальні інформаційні системи підтримки прийняття рішень, які можуть адаптуватися до різного набору критеріїв та їх пріоритетів. В основі таких систем може бути реалізація одного з відомих експертних методів, а також використання методів багатокритеріальної оптимізації.

Іншим підходом може бути натренована штучна нейронна мережа, яка дасть бажаний результат з припустимою похибкою за досить короткий час, проте постає питання про створення коректного і достатнього набору даних для її тренування. Варто зазначити, що подібна мережа може бути використана як у складі експертної системи для більшого ступеня синергії з іншими компонентами, так і окремо.

У публікаціях [11-14] науковці дослідили можливість застосування симуляції для

оцінки діджиталізації. Зокрема, у [14] використано приклад підприємства, що займається виробництвом сталі. В рамках експерименту вони розробили загальний сценарій цифровізації, описали процес діджиталізації та оцінки ризиків, а також розробили детальну модель, що дозволила симулювати безпосередньо діджиталізацію в рамках обраних виробничих процесів. Такий підхід має перевагу за рахунок того, що можна достатньо точно оцінити можливий вплив на бізнес-процеси від впроваджених заходів без будь-якого втручання у виробництво або надання послуг.

Інтелектуальні системи як основа оцінювання ефективності діджиталізації. Інтелектуальні системи являють собою комп'ютерну систему, що здатна взаємодіяти з навколишнім середовищем. Часто такі системи використовуються для вирішення задач прийняття рішень. Зазвичай інтелектуальні системи складаються з бази знань, інтерфейсу взаємодії, а також механізму пошуку рішень. Механізм пошуку рішень може бути як, наприклад, звичайною математичною моделлю, так і складним компонентом з елементами штучного інтелекту.

Для задач оцінки ефективності діджиталізації доцільно використовувати інформаційні інтелектуальні системи, що є комплексом, який поєднує як програмно-технічні та математичні засоби, так і лінгвістичні. Прикладом

такої системи може бути експертна система. Така система може підібрати необхідний метод оцінки на основі набору вхідних даних (набір факторів, статистичні дані тощо), тому таким чином відбуватиметься симуляція процесу прийняття рішення експертом-людиною.

Елементи побудови автоматизованих систем оцінки ефективності діджиталізації. Спочатку доцільно дати визначення автоматизованої системи й автоматизованої інформаційної системи. Згідно з ДСТУ [15] автоматизована система – це організаційно-технічна система, що складається із засобів автоматизації певного виду чи кількох видів діяльності людей та персоналу, що здійснює цю діяльність. В свою чергу, Закон України «Про захист інформації в інформаційно-телекомунікаційних системах» наводить визначення автоматизованої інформаційної системи [16]: це організаційно-технічна система, в якій реалізується технологія обробки інформації з використанням технічних і програмних засобів. Якщо застосувати подані визначення до оцінки ефективності діджиталізації, то подібна система повинна автоматизувати безпосередньо процес оцінки доцільності впроваджених заходів за певним набором критеріїв.

Якщо скомбінувати особливості інтелектуальної й автоматизованої інформаційної систем, то отриманий результат можна зобразити у вигляді схеми, наведеної на рисунку 1.

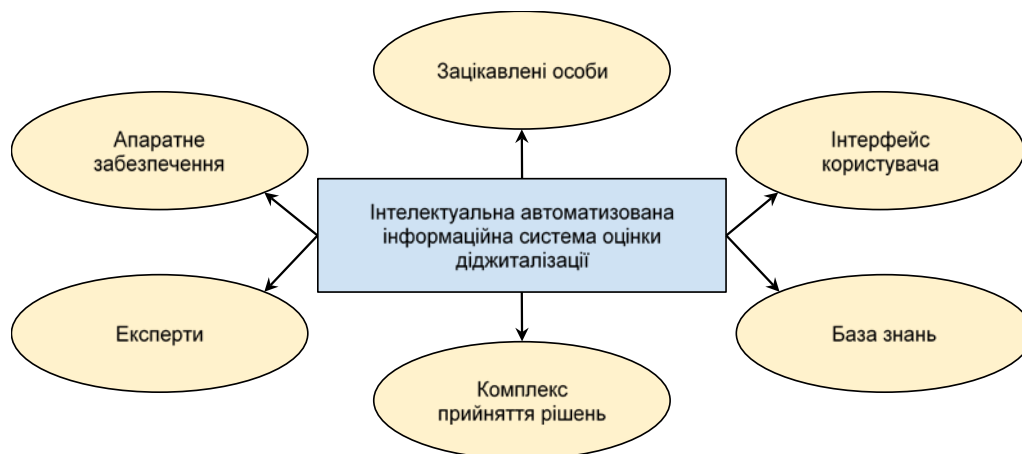


Рисунок 1 – Елементи інформаційної системи оцінки ефективності діджиталізації

Джерело: побудовано на основні типових структур інтелектуальних систем

Детальніше розберемо кожен зі складових.

Зацікавлені особи є представниками підприємств чи бізнесів, де планується впроваджувати або вже впроваджується діджиталі-

зація. За допомогою *інтерфейсу користувача* вони вносять вхідні дані до системи, а також роблять запити на оцінку цифровізації.

Апаратне забезпечення – це фізична інфраструктура, що призначена для виконання

обчислювальної роботи, зберігання даних тощо. До прикладів апаратного забезпечення можна віднести комп'ютери, сервери (в т. ч. хмарні), портативні пристрої тощо.

Експерти беруть участь у розробці критеріїв оцінки, їх ваги, пріоритету, групування тощо в контексті конкретної предметної області. Вся вищезазначена інформація зберігається у *базі знань*. Також вони допомагають у створенні алгоритмів аналізу статистичних даних. *Експерти* взаємодіють із системою, використовуючи *інтерфейс користувача*.

Комплекс прийняття рішень – це сукупність математичних, лінгвістичних, технічних засобів, що роблять можливим процес прийняття рішень. Сюди також можуть входити компоненти з елементами штучного інтелекту. Цей комплекс можна розглядати як підсистему основної системи. Він отримує вхідні дані та алгоритми оцінки із бази знань, а як результат дає розгорнутий аналіз впроваджених заходів.

База знань являє собою сховище даних для всього набору статистичних даних, інформації про критерії та принципів логіки оцінки бізнес-процесів.

В свою чергу, *інтерфейс користувача* використовується для взаємодії із системою *експертами* та *зацікавленими особами*. По своїй суті – це комп'ютерна програма, яка може бути виконана у вигляді, наприклад, Web-додатка, настільного додатка чи мобільного застосунку. У кожного варіанта виконання є свої переваги та недоліки, проте найбільш універсальним підходом був би Web-додаток, оскільки його можна використовувати на будь-якому пристрої, що має доступ до мережі Інтернет. *Інтерфейс користувача* взаємодіє з *комплексом прийняття рішень* для обробки запитів на оцінку та *базою знань* для збереження вхідних даних, необхідних для функціонування системи.

Інтерфейс користувача, база знань та комплекс прийняття рішень використовують апаратне забезпечення для своєї роботи.

Результати досліджень особливостей, проблем та задач впровадження діджиталізації свідчать, що діджиталізація бізнесу як еволюційний процес впровадження знань і творчості є невід'ємною частиною успішних компаній. Діджиталізація дозволяє оптимізувати бізнес-процеси, підвищити прибутковість компанії та орієнтована на створення цінності для задоволення потреб споживачів.

У процесі впровадження діджиталізації та оцінки її ефективності можна застосувати інтелектуальні інформаційні системи підтримки прийняття рішень, які можуть адаптуватися до різного набору критеріїв та їх пріоритетів. В основі таких систем може бути реалізація одного з відомих експертних методів, а також використання методів багатокритеріальної оптимізації.

До елементів інформаційної системи оцінки ефективності діджиталізації можуть належати: зацікавлені особи, апаратне забезпечення, експерти, комплекс прийняття рішень, база знань, інтерфейс користувача.

Обговорення результатів дослідження. Отримані результати показали, що тема впровадження діджиталізації є достатньо актуальною, а спроби оцінити результативність цього процесу знаходять відображення в роботах багатьох науковців.

Основні проблеми впровадження діджиталізації пов'язані з визначенням коректних або ефективних критеріїв оцінки впроваджених заходів, ретроспективою невдалих спроб діджиталізації, нехтуванням кібербезпекою або приділенням їй недостатньої уваги.

Висновки. Виконаний аналіз поточного стану технологій та методів оцінки ефективності діджиталізації свідчить, що ефективність впровадження діджиталізації в різних сферах людської діяльності оцінюється за допомогою методів багатокритеріальної оптимізації, які використовуються для пошуку значень розроблених індексів, як-то метод зважених сум. Нерідко використовуються експертні опитування як інструмент для прийняття рішень.

Наукова новизна проведеного дослідження полягає у визначенні особливостей процесу діджиталізації у контексті оцінювання її ефективності. Оптимізація бізнесу шляхом впровадження ІТ-рішень та роботизації включає нові зовнішні фактори та можливості, які потребують перегляду стратегічних пріоритетів.

Практичне значення отриманих результатів полягає у визначенні структури інформаційної системи оцінки ефективності діджиталізації, яка може бути використана для розробки реальної системи оцінювання ефективності цифрової трансформації.

Подальшого дослідження потребують питання визначення особливостей проектування та планування процесу діджиталізації,

питання структуризації та застосування критеріїв оцінювання ефективності цифрової трансформації, а також управління процесом розробки інтелектуальної інформаційної системи оцінювання ефективності діджиталізації.

Список використаних джерел

- [1] M. Bogavac, and Z. Cekerevac, "IDSME index – new method for evaluation of SMEs digitalization", *MEST Journal*, vol. 2, no. 7, pp. 9-20, 2019.
- [2] European Commission, Digital Economy and Society Index – Methodology Note [Online]. Available: https://ec.europa.eu/information_society/newsroom/image/document/2018-20/desi-2018-methodology_E886EDCA-B32A-AEFB-07F5911DE975477B_52297.pdf. Accessed on: March 18, 2022.
- [3] A. Valiyev, F. V. oglu Rustamov, R. A. Huseynova, M. S. Orujova, and S. N. Musayeva, "The digitalization effectiveness as an innovative factor development of the agriculture in Azerbaijan", *JEECAR*, vol. 9, no. 2, pp. 194-205, 2022.
- [4] O. Pakhnenko, P. Rubanov, D. Hacar, V. Yatsenko, and I. Vida, "Digitalization of financial services in European countries: Evaluation and comparative analysis", *Journal of International Studies*, no. 2, pp. 267-282, 2021.
- [5] V. Paulauskas, L. Filina-Dawidowicz, and D. Paulauskas, "Ports digitalization level evaluation", *Sensors*, vol. 21, no. 18, p. 6134, 2021.
- [6] A. Geyda, "Models and methods to estimate digitalization success predictively", in *Proc. 21st Int. Workshop on Computer Science and Information Technologies*, 2019.
- [7] Т. В. Лазоренко, та І. Л. Шолом, "Діджиталізація як основний фактор розвитку бізнесу", на *I Міжнар. наук.-практ. конф. Бізнес, інновації, менеджмент: проблеми та перспективи*, Київ, 2020, с. 50-51.
- [8] Український інститут майбутнього. "Україна 2030Е – країна з розвинутою цифровою економікою". [Електронний ресурс]. Режим доступу: <https://strategy.uifuture.org/kraina-z-rozvinutoyu-cifrovoyu-ekonomikoyu.html>. Дата звернення: March 18, 2022.
- [9] Common Digital Transformation Challenges & How to Overcome Them [Online]. Available: <https://www.panorama-consulting.com/>

digital-transformation-challenges. Accessed on: June 12, 2021.

- [10] Digitalization Problems [Online]. Available: <https://gotodigital.es/en/digitalization-problems>. Accessed on: June 12, 2021.
- [11] F. Misrudin, and Lee Ching Foong. "Digitalization in semiconductor manufacturing-simulation forecaster approach in managing manufacturing line performance", *Procedia Manufacturing*, vol. 38, pp. 1330-1337, 2019.
- [12] R. Aglave, J. Lusty, and J. Nixon, "Using simulation and digitalization for modular process intensification", *Chemical Engineering Progress*, vol. 3, pp. 45-49, 2019.
- [13] L. D. Collou, G. W. J. Bruinsma, and M. J. van Riemsdijk, "Digitalization of HRM: Designing a simulation model for HR decision making: The development of a simulation for the alignment of HR-practices", in *Proc. Int. Conf. of the Dutch HRM Network*, Tilburg, Netherlands, 2019.
- [14] F. Schlüter, E. Hettterscheid, and M. Henke, "A simulation-based evaluation approach for digitalization scenarios in smart supply chain risk management", *Journal of Industrial Engineering and Management Science*, vol. 2017, no. 1, pp. 179-206, 2019.
- [15] ДСТУ 2960-94 Організація промислового виробництва. Основні поняття. Терміни та визначення.
- [16] Верховна Рада України (1994, лип. 05), *Закон України № 80/94-ВР, Про захист інформації в інформаційно-телекомунікаційних системах*, *Відомості Верховної Ради України*, № 31, ст. 286, 1994. [Електронний ресурс]. Режим доступу: <https://zakon.rada.gov.ua/laws/show/81/94-%D0%B2%D1%80>.

References

- [1] M. Bogavac, and Z. Cekerevac, "IDSME index – new method for evaluation of SMEs digitalization", *MEST Journal*, vol. 2, no. 7, pp. 9-20, 2019.
- [2] European Commission, Digital Economy and Society Index – Methodology Note [Online]. Available: https://ec.europa.eu/information_society/newsroom/image/document/2018-20/desi-2018-methodology_E886EDCA-B32A-AEFB-07F5911DE975477B_52297.pdf. Accessed on: March 18, 2022.

- [3] A. Valiyev, F. V. oglu Rustamov, R. A. Huseynova, M. S. Orujova, and S. N. Musayeva, "The digitalization effectiveness as an innovative factor development of the agriculture in Azerbaijan", *JEECAR*, vol. 9, no. 2, pp. 194-205, 2022.
- [4] O. Pakhnenko, P. Rubanov, D. Hacar, V. Yatsenko, and I. Vida, "Digitalization of financial services in European countries: Evaluation and comparative analysis", *Journal of International Studies*, no. 2, pp. 267-282, 2021.
- [5] V. Paulauskas, L. Filina-Dawidowicz, and D. Paulauskas, "Ports digitalization level evaluation", *Sensors*, vol. 21, no. 18, p. 6134, 2021.
- [6] A. Geyda, "Models and methods to estimate digitalization success predictively", in *Proc. 21st Int. Workshop on Computer Science and Information Technologies*, 2019.
- [7] T. V. Lazorenko, and I. L. Sholom, "Digitalization as a major factor in business development", in *Proc. I Int. Sci. and Pract. Conf. Business, Innovations, Management: Problems and Perspectives*, Kyiv, 2020, pp. 50-51 [in Ukrainian].
- [8] Ukrainian Institute of the Future. "Ukraine 2030E is a country with a developed digital economy". [Online]. Available: <https://strategy.uifuture.org/kraina-z-rozvinutoyu-cifrovoyu-ekonomikoyu.html> Accessed on: Febr. 12, 2022 [in Ukrainian].
- [9] Common Digital Transformation Challenges & How to Overcome Them [Online]. Available: <https://www.panorama-consulting.com/digital-transformation-challenges>. Accessed on: June 12, 2021.
- [10] Digitalization Problems [Online]. Available: <https://gotodigital.es/en/digitalization-problems>. Accessed on: June 12, 2021.
- [11] F. Misrudin, and Lee Ching Foong. "Digitalization in semiconductor manufacturing-simulation forecaster approach in managing manufacturing line performance", *Procedia Manufacturing*, vol. 38, pp. 1330-1337, 2019.
- [12] R. Aglave, J. Lusty, and J. Nixon, "Using simulation and digitalization for modular process intensification", *Chemical Engineering Progress*, vol. 3, pp. 45-49, 2019.
- [13] L. D. Collou, G. W. J. Bruinsma, and M. J. van Riemsdijk, "Digitalization of HRM: Designing a simulation model for HR decision making: The development of a simulation for the alignment of HR-practices", in *Proc. Int. Conf. of the Dutch HRM Network*, Tilburg, Netherlands, 2019.
- [14] F. Schlüter, E. Hetterscheid, and M. Henke, "A simulation-based evaluation approach for digitalization scenarios in smart supply chain risk management", *Journal of Industrial Engineering and Management Science*, vol. 2017, no. 1, pp. 179-206, 2019.
- [15] DSTU 2960-94 Organization of industrial production. Basic concepts. Terms and definitions [in Ukrainian].
- [16] Verkhovna Rada of Ukraine (1994, July 05), *The Law of Ukraine no. 80/94-VR, On information protection in information and telecommunication systems*, *Vidomosti Verkhovnoi Rady Ukrainy*, no. 31, art. 286, 1994. [Online]. Available: <https://zakon.rada.gov.ua/laws/show/81/94-%D0%B2%D1%80> [in Ukrainian].

I. H. Zinchenko, Ph. D. Student,

e-mail: igor.zinchenko557@gmail.com

O. V. Lavdanska, Ph. D., Associate Professor,

Cherkasy State Technological University

Shevchenko blvd, 460, Cherkasy, 18006, Ukraine

MODERN TECHNOLOGIES FOR EVALUATING THE EFFECTIVENESS OF DIGITALIZATION

The paper examines modern technologies for evaluating the effectiveness of digitalization. The process of digitalization is analyzed. It is determined that an integral part of successful companies is a continuous rethinking of the company's development method aimed at a simple, economical and reliable work option, updating and following new trends in the market of technical developments. It has been established that digital transformation allows to optimize business processes, increase the profitability of the company and is focused on creating value to meet the needs of consumers. The stages of

the digitalization implementation project, which include: analysis of current resources and processes; development of a plan for the transition to the use of new technologies together with the final goals; implementation of planned activities; evaluation and systematization of the obtained results, are presented. The main tasks of digitalization, such as: development and implementation of a plan for the transition to a new paradigm; migration of information flows into digital form; reduction or elimination of the human factor in typical tasks; automation of processes that can be automated, are indicated. The main problems of digitalization implementation, namely: determination of incorrect or ineffective criteria for evaluation of implemented measures; setting too abstract goals in the approach to the digitalization process; necessity to think ahead and build a map of interaction with the client, effective for both interested parties; retrospective of unsuccessful attempts at transformation; neglect or insufficient attention to cyber security, are presented. An analysis of the current state of technologies and methods for assessing the effectiveness of digitalization is performed. It is determined that in the process of implementing digitalization and evaluating its effectiveness, it is possible to apply intelligent decision support information systems that can adapt to a different set of criteria and their priorities. The basis of such systems may include the implementation of one of the well-known expert methods, as well as the use of multi-criteria optimization methods. The elements of the information system for evaluating the effectiveness of digitalization may include: interested persons, hardware, experts, decision-making complex, knowledge base, user interface. In further research, it is advisable to create a mathematical model for evaluating efficiency criteria and a model for evaluating the digitalization process itself.

Keywords: digitalization, evaluation, efficiency, system, automation, decision-making.

D. O. Yakymenko, Graduate Student (Applicant),
Ye. Yu. Kataieva, Ph. D., Associate Professor
Cherkasy State Technological University
Shevchenko blvd, 460, Cherkasy, 18006, Ukraine

METHODS AND MEANS OF INTELLIGENT ANALYSIS OF TEXT DOCUMENTS

The paper reviews the methods of analysis and processing of electronic documents. Methods of analysis of text documents to solve the problem of determining the thematic affinity of texts are analyzed. An overview of existing approaches to solving the classification problem is performed. The main approaches used in the task of text classification are described; the stages of the classification process are determined and the most common methods of classifying text documents are considered. The main approaches to text pre-processing, such as: lower case, root correction, stemming, lemmatization, stop word removal, normalization, are considered. Advantages and disadvantages of each approach are considered. The procedure for reducing the dimension of a set of features with a division into sub-processes: selecting features and highlighting features is considered.

Keywords: keywords, text analysis, search, text documents, classifications.

Introduction. The rapid development of computer technology has led to the rapid accumulation of electronic text documents. This situation will lead to the fact that existing methods of processing electronic text documents will not be able to meet the needs of users in both corporate networks and the Internet.

That is why there is a need for methods that will provide a quick and easy distribution of documents by category or keyword.

Based on this, you can identify the following major problems associated with increasing the amount of information:

- the growing volume of documents posted on the Internet is the reason for the growing difficulties in finding the necessary documents for users and their organization;

- technologies for working with text documents involve a rather large complexity of implementation, which affects the speed of data processing;

- a significant part of electronic text documents is unstructured information.

The purpose of the study is to analyze existing means of electronic texts analysis.

The research task consists in the analysis of:

- 1) means to work with text documents;
- 2) methods of processing text documents;
- 3) applications used to analyze electronic texts.

Presentation of the main material. The most conventional examples of intelligent text analysis include technologies for extracting

factual information about the search object, fuzzy search; thematic and tonal (accurate and complete) rubrics; selection of documents; selection of key topics; construction of annotations; use of methods of intelligent text analysis to determine the directions of research of large document funds and obtain new information about the process of automatic text summarization. The most modern areas of obtaining information from texts today are:

- analytical processing of facts;
- file keeping;
- annotation of documents;
- conducting a thematic analysis of docskills (clustering and heading);
- construction and dynamic analysis of a sample of the ethical structure of texts;
- selection of key topics and information of other objects;
- study of frequency characteristics of texts [1].

Most of these methods are aimed at processing documents, each of which relates to a specific, often quite narrow topic. This statement is true mainly for small documents (such as small web pages), but not for large documents, which often relate to several topics at once. Examples of such documents are musts, scientific articles and others. We introduce a definition for this type of document.

Definition 1. A text document of any size, which simultaneously belongs to several topics, is called polythematic.

Definition 2. A text document of any size, which relates to only one specific topic, is called monothematic.

Definition 3: The subject of the text document - some subjective perception of the person, the user of the search engine, which is considered in the text of the subject area, its main content.

Definition 4: A text is a sequence of sentences, words, built according to the rules of a given language, a given sign system and forms a message that carries some useful information [2].

Polythematic documents contain multifaceted goals, thus reflecting the information needs of different users. Text repositories, which usually store such documents, are usually characterized by several topics.

To date, few methods allow to process polythematic text documents - to perform their classification or clustering. Therefore, the development of methods for solving these problems is especially important. The purpose of the classification of polythematic text documents is to assign the same text document to more than one topic. The purpose of clustering polythematic text documents is to automatically divide text documents into a priori not specified groups so that each document belongs to more than one of them [3, 4].

Classification of text documents

Document classification, often known as document categorization, is a challenge in library, information, and computer science. The assignment of a document to one or more groups or categories is the task. This can be accomplished "manually" or algorithmically. Intelligent document categorization has traditionally been the domain of library science, whereas algorithmic document classification has traditionally been the domain of informatics and computer science.

Texts, photos, music, and other media may all be categorized. Each type of document has its own set of categorization issues. Text categorization is assumed unless otherwise stated.

The tasks of automatic document classification can be divided such types: supervised document classification where some external mechanism (such as human feedback) provides information on the correct classification of documents, unsupervised document classification (also known as document clustering), semi-supervised document classification, in which sections of the documents are labelled by an external process, and unsupervised document clas-

sification, in which the classification must be performed totally without reference to external information. There are different software programmes available with varied licence models.

Documents can be categorized based on their subjects or other attributes (such as document type, author, year of publication, etc.). Only sub-subject categorization is discussed in the remainder of this article. There are two basic approaches to document subject classification: the content-based approach and the request-based approach. The categorization task is as follows: there are several papers and classifications. The aim is to discover pairs of "document, category" that are compatible with each other [5].

The general classification scheme consists of four stages:

- 1) pre-processing and indexing;
- 2) reducing the dimensionality of many features;
- 3) construction and training of the classifier;
- 4) assessment of the quality of classification.

At the stage of preliminary processing and indexing of documents, the features of the document, which are all its significant words or phrases, are formed. This stage includes tokenization, i.e. breaking the text into smaller objects, such as sentences, phrases or words; removal of functional words (semantically neutral, such as conjunctions, prepositions, articles, etc.) and morphological analysis, i.e. identification of parts of speech and stemmatization or lemmatization.

Reducing the dimensionality of a set of features is a process of giving weight to words, depending on their importance for the classification of the text, and further removing weightless terms from the set of features.

A threshold weight is set to remove terms below which terms are considered unimportant.

By reducing the dimensionality of the set of terms, you can reduce the effect of retraining - a phenomenon in which the classifier focuses on random or erroneous characteristics of educational data, and not on actually important ones [6].

We are only interested in the first and second stages. Therefore, let's consider them in more detail.

Pre-processing of text documents

Pre-processing your text simply means transforming it into a predictable and analyzable format for your work. A task is a mix of method and domain in this context. A Task may be ex-

tracting top keywords from Tweets (domain) using tfidf (method).

There are several methods for text pre-processing. Here are various techniques that you should be aware of, and we'll attempt to emphasize the significance of each.

Lowercasing – although it is sometimes forgotten, lowercasing of all text data is one of the simplest and most efficient forms of text preparation. It is applicable to the majority of text mining tasks and might be useful when your dataset is small. It also considerably improves the consistency of the predicted output [7].

Stemming – the practice of reducing word inflections (for example, troubled, troubles) to their base form is known as stemming (e.g., trouble). In this situation, the "root" might be a canonical variant of the original word rather than a true root word.

Stemming is a primitive heuristic procedure that cuts off the ends of words in the goal of appropriately changing them into their base form.

Lemmatization – on the surface, lemmatization appears to be similar to stemming in that the objective is to remove inflections and map a word to its root form. The main distinction is that lemmatization attempts to do it correctly. It doesn't merely cut things off; it truly changes the root of the term.

Stop words removal – language stop words are a group of frequently used terms. Stop words in English include "a," "the," "is," "are," and others. Stop words are used with the theory that by eliminating uninformative terms from the text, we can focus on key words instead.

Stop words removal was proved to be ineffective in classification systems, although being successful in search and topic extraction systems. However, it helps to reduce the number of features considered, keeping your models reasonably sized [8].

Normalization – text normalization is an often-overlooked pre-processing procedure. The process of converting a text into a canonical (standard) form is known as text normalization.

Text normalization has also been shown to be useful in evaluating highly unstructured clinical texts in which physicians take notes in unconventional ways. It's particularly effective for subject extraction if there are a lot of close synonyms and spelling discrepancies.

Unlike stemming and lemmatization, there is no universal method for normalizing texts. It is usually determined by the task.

Dictionary mappings (the easiest), statistical machine translation (SMT), and spelling-correction-based techniques are some typical ways to text normalization.

The optimal preparation for one job may become the greatest nightmare for another. Take note: text preparation cannot be transferred straight from job to task. Currently, there are two basic ways to text document pre-processing, known as document image generation.

1. Statistical approach. The methods of this approach, as a rule, consist in the analysis of the frequency of occurrence of words in texts in one or another of its variation and in the use of this information in the course of revealing and selecting representative signs of documents.

2. Linguistic approach. The main directions of this approach are morphological and syntactic analysis. Morphological analysis is important for Ukrainian language documents, as it allows to bring the processed features to some normal form. Such a reduction is necessary to recognize equivalent words with a common morphological basis. Parsing allows to automatically parse a text and build syntactic structures of its phrases, grouping words into classes within which they have similar syntactic behavior, these classes of words are called syntactic or grammatical categories (for example, parts of speech).

Linguistic and statistical methods complement each other because they use different approaches to information analysis, so the best results can potentially be achieved by combining both methods. However, currently linguistic methods (except for morphological analysis) are usually used only for research purposes and are not used in ready-made (commercial) information retrieval systems, as this approach is associated with much higher computational costs than statistical one. Moreover, a significant improvement in the quality of such systems in this approach compared to the statistical approach hasn't been found, which is often due to the large noise component introduced by the errors of the automated parser. Therefore, the most common today is the combined method of pre-processing of text documents, which combines statistical approach and morphological analysis. This is a compromise between the complexity of calculations and the quality of results [9].

Since the linguistic approach to pre-processing of text documents is currently practically not used, as it requires large computational resources, we consider static methods that do not

require such large computing resources, and are common methods of pre-processing of text documents.

In the statistical approach, the pre-processing of text documents is to perform the following main steps:

- 1) formation of the space of documents signs;
- 2) display of images docskills in the space of their signs;
- 3) reduction of the initial avsharp signs of documents [10].

Word2vec is a technique for natural language processing published in 2013. The word2vec algorithm uses a neural network model to learn word associations from a large corpus of text. Once trained, such a model can detect synonymous words or suggest additional words for a partial sentence. As the name implies, word2vec represents each distinct word with a particular list of numbers called a vector. The vectors are chosen carefully to indicate the level of semantic similarity between the words represented by those vectors [11].

Word2vec is a group of related models that are used to produce word embeddings. These models are shallow, two-layer neural networks that are trained to reconstruct linguistic contexts of words. Word2vec takes as its input a large corpus of text and creates a vector space, typically of several hundred dimensions, with each unique word in the corpus being assigned a corresponding vector in the space. Word vectors are positioned in a vector space such that words that share common contexts in the corpus are located close to each other in the space [12].

Analysis of methods of text data processing

The method of determining the weights of the features of the document. The most common methods for determining the weights of document features are TF-IDF, a statistical indicator used to assess the importance of words in the context of a document that is part of a collection of documents or corpus. The weight (significance) of a word is proportional to the number of uses of that word in the document, and inversely proportional to the frequency of use of the word in other documents of the collection.

The TF-IDF indicator is often used to represent the documents of a collection as numerical vectors that reflect the importance of using each word from a set of words (the number of words in the set determines the dimension of the vector)

in each document. Such a model is called a vector model and makes it possible to compare texts by comparing their representing vectors in a certain metric (Euclidean distance, cosine measure, Manhattan distance, Chebyshev distance, etc.), perform cluster analysis [13].

The first step in word processing is to calculate the TF-IDF weights for each word ω in each document. Φ_p^* .

TF is the ratio of the number of occurrences of the selected word to the total number of words in the document. Thus, the importance of the word within the selected document is assessed [14]:

$$tf(\omega, \Phi_p^*) = \frac{n_{\omega\Phi_p^*}}{n_{\Phi_p^*}}, \quad (1)$$

where $n_{\omega\Phi_p^*}$ is the number of occurrences of the word ω in the document;

$n_{\Phi_p^*}$ is the total number of words in the document.

IDF (Inverse Document Frequency) is the inversion of the frequency with which a word occurs in the documents of a collection. The use of IDF reduces the weight of commonly used words [14]

$$idf(\omega, A) = \frac{|\Phi_p^*|}{|\Phi_p^* \supset \omega|}, \quad (2)$$

where $|\Phi_p^*|$ is the total number of documents in a collection A;

$|\Phi_p^* \supset \omega|$ is the number of documents in which it occurs when $n_{\omega\Phi_p^*} \neq 0$.

So, the TF-IDF indicator is [14]

$$tfidf(\omega, \Phi_p^*, A) = tf(\omega, \Phi_p^*) * idf(\omega, A). \quad (3)$$

More weight of TF-IDF will be given to words with a high frequency of occurrence within the document and low frequency of use in other documents of the collection [14].

The method of assessing the proximity of text documents. Polythematic text documents can also be characterized by using document similarity features.

One of the methods to calculate relative estimates of thematic similarity includes the following steps:

- for each document, several (relatively small) documents representing its thematic environment are defined;
- constructed thematic environments are analyzed in order to form a set of key themes, which characterize the subject of the source doc-

ument in relation to other documents of the collection;

- obtained keyword sets are used to further calculate relative estimates of thematic similarity.

The need to find the thematic environment of a document is caused by the relativity of the concept of thematic proximity of documents, which is determined by the context in which the proximity of documents is assessed. For example, two documents characterizing, respectively, changes in stock quotes and changes in exchange rates, are likely to be considered thematically similar among a random set of documents, but at the same time they differ significantly within the limits of a highly specialized economic collection. Therefore, the assessment of thematic proximity is not only determined by the documents themselves, but also depends on the whole array of documents [15].

It is known that vocabulary and frequency of use of words depend on the subject. Therefore, only those words that are more specific to the subject of this document are taken into account to calculate estimates of thematic proximity. Such words are distinguished by the results of the analysis of the approximated thematic environment of this document.

Keyword selection method. Using the method of extracting keywords from the text allows to find the information you need in a short period of time.

Keyword is a word or constant expression of natural language, which is used to express some aspect of the content of the document; a word that has a significant semantic load. It can be a key when searching for information on the Internet or on a website.

There may be a synonymous relationship between keywords in terms of this search engine. The accumulation of keywords through meaningful analysis of texts or algorithmically, for example, by comparing words of a text with a fixed list of non-keywords, is an important step in choosing the source dictionary of information retrieval languages; the selected keywords are further combined into descriptors. Descriptor dictionaries provide links from keywords to corresponding descriptors. Keywords are the basis of search results. It should be remembered that a keyword can be not only a word but also a phrase [16].

Selection of keywords, extraction of the most important or characteristic fragments from one or many sources of information has become

an integral part of our lives. Keyword selection tools are certainly useful, but their capabilities are limited to selecting and extracting the original snippets from the original document and combining them into a short text. The preparation of a summary aims to describe the main content of the text.

The main difference between the means of selecting keywords is that they, in fact, form a summary or set of citations from a particular material. Both types of presentation have two main purposes: to identify the most important idea of the full text and select keywords [17].

You can identify three key points that are not taken into account when selecting keywords from the text:

- a) division into parts by keywords (by formatting);
- b) selection of keywords (by weight);
- c) formation of a summary document (in the form of statistics).

The method of comparison is the comparison of objects in order to identify common features or differences between them. The method of comparison is used in the process of generalization, when it is necessary to identify identities, coincidences and contradictions in the objects of study. Here identity is a full-fledged coincidence of all signs; coincidence is a coincidence of signs, starting from one; contradiction is when features of some objects are absent in others.

One of the methods used to identify clusters of documents that have similar properties only in some aspects, such as words or images, is *biclusterization*. The method is used for queries and indexing of full-text systems. Initial data is a matrix in which rows correspond to words and columns correspond to documents. To cluster documents, the number of word occurrences in a document, the total number of documents, and the number of documents containing a particular word are taken into account [18]. Thus, words can be clustered based on the documents in which they occur. Clusters are convenient for automatic construction of statistical thesauri, clarification of queries and automatic classification of documents, but it is impossible to perform meaningful text analysis using clusters.

Search for fuzzy duplicates involves clustering documents by the similarity of their certain characteristics, and the implementation of the algorithm consists of the following steps:

- canonization of steps – at this stage the text is cleaned of unnecessary words that do not

carry a meaning during comparison, i.e. the text is reduced to a canonical form;

- breaking the text into shingles;
- finding checksums – unique numbers,

each of which corresponds to a text and the function of its calculation. Then, from the whole set of checksums (their number is equal to the number of words in the document minus ($\hat{W}-1$, where EP is the number of words in the shingles), only those that are divisible by a certain selected premature number are selected;

- search for identical sequences - one shingle, which coincided during the selection, approximately corresponds to a predetermined number of identical parts in the full text [19].

Lexical and syntactic templates are characteristic expressions (phrases or inversions), constructions from the corresponding elements of natural language. Such templates allow to build a semantic model of the text. It is assumed that lexical relations in the document can be described with the help of templates (samples). This method uses a hierarchy of templates consisting of indicators of parts of speech and group symbols.

Thus, the analysis of the software used to detect text documents has shown the imperfection of the methods underlying modern automatic analysis systems. So, there is a need to create such software that would use the methods of deep linguistic processing. This will make it possible to compare text documents by content [20].

Analysis of word processing software

OBSERVER. This system offers an approach to using existing ontologies to access distributed and independently developed data repositories. It is assumed that there are many pre-created ontologies of subject areas, and the user does not have to "adapt" to a particular ontology. The user formulates his language query in terms of one or more ontologies, and the broker "searches" for relevant documents by translating the query into suitable ontologies and, if necessary, combining several ontologies to respond more accurately to the query [21].

TEXTANALYST. TextAnalyst is designed as a tool for analyzing the content of texts, semantic search for information, the formation of electronic archives, and provides the user with the following main features:

- 1) analysis of the content of the text with automatic formation of a semantic network with hyperlinks - obtaining a semantic portrait of the text in terms of basic concepts and their semantic connections;

- 2) analysis of the content of the text with automatic formation of the thematic tree with hyperlinks - identification of the semantic structure of the text in the form of a hierarchy of topics and subtopics;

- 3) semantic search, taking into account the hidden semantic connections of the query words with the words of the text;

- 4) automatic abstracting of the text - the formation of its semantic portrait in terms of the most informative phrases;

- 5) clustering of information - analysis of the distribution of text material by thematic classes;

- 6) automatic indexing of text with conversion to hypertext;

- 7) ranking of all types of information about the semantics of the text by "degrees of significance" with the possibility of varying the detail of its research [22].

ADVEGO PLAGIATUS - software product that allows to perform semantic analysis online, while finding the semantic core, and also allows to check the electronic document for uniqueness against web documents that are publicly available on the Internet.

The text of the article is analyzed by several algorithms:

- Shingles algorithm - exact matches of phrases, cypypaste sources, as well as pages on which stolen texts are placed are checked.

- Algorithm of lexical matches - checks the similarity of a set of lemmas, terms and significant words, there are sources of rewriting, as well as pages that match the topic with the article being tested.

- Pseudo-unification algorithm - the presence of third-party characters and signs of text processing is checked by "enhancing" uniqueness services [23].

The main task of these systems is to search for information in large full-text arrays. The database of such systems can download any textual sources of information, including large ones: encyclopedias, directories, archives of periodicals, entire libraries of special literature, archives of corporate documents, specialized archives, such as historical, patent, court, transcripts, protocols and much more. In response to a specific request, the system issues a set of links. Next, the system must process each link and publish all relevant texts, i.e. the system must search not just documents, but the information contained in them.

DEEPDIVE is a system to extract value from dark data. Like dark matter, dark data is the great mass of data buried in text, tables, figures, and images, which lacks structure and so is essentially unprocessable by existing software. DeepDive helps bring dark data to light by creating structured data (SQL tables) from unstructured information (text documents) and integrating such data into an existing structured database. DeepDive is used to extract sophisticated relationships between entities and make inferences about facts involving those entities. DeepDive helps to process a wide variety of dark data and put the results into a database. With the data in a database, one can use a variety of standard tools that consume structured data.

DeepDive is a trained system that uses machine learning to deal with various forms of noise and imprecision. DeepDive is designed to make it easier for users to train the system through low-level feedback and rich, structured domain knowledge via rules. DeepDive wants to help experts who do not have machine learning expertise. One of DeepDive's key technical innovations is its ability to solve statistical inference problems at massive scale.

DeepDive differs from traditional systems in several ways:

- DeepDive asks the developer to think about features - not algorithms. In contrast, other machine learning systems require the developer to think about which clustering algorithm, which classification algorithm, etc. to use.
- DeepDive systems can achieve high quality: DeepDive has higher quality than human volunteers in extracting complex knowledge in scientific domains and winning performance in entity relation extraction competitions.
- DeepDive is aware that data is often noisy and imprecise: names are misspelled, natural language is ambiguous, and humans make mistakes. Taking such imprecision into account, DeepDive computes calibrated probabilities for every assertion it makes.
- DeepDive is able to use large amounts of data from a variety of sources. Applications built using DeepDive have extracted data from millions of documents, web pages, PDFs, tables, and figures.
- DeepDive allows developers to use their knowledge of a given domain to improve the

quality of the results by writing simple rules that inform the inference (learning) process. DeepDive can also take into account user feedback on the correctness of the predictions to improve the predictions.

- DeepDive is able to use the data to learn "distantly". In contrast, most machine learning systems require tedious training for each prediction. In fact, many DeepDive applications, especially in early stages, need no traditional training data at all!

- DeepDive's secret is a scalable, high-performance inference and learning engine. For the past few years, we have been working to make the underlying algorithms run as fast as possible [24].

SCIKIT-LEARN (formerly `scikits.learn` and also known as `sklearn`) is a free software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms, including support-vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interact with the Python numerical and scientific libraries NumPy and SciPy. Scikit-learn is a NumFOCUS fiscally sponsored project.

Scikit-learn is largely written in Python, and uses NumPy extensively for high-performance linear algebra and array operations. Furthermore, some core algorithms are written in Cython to improve performance. Support vector machines are implemented by a Cython wrapper around LIBSVM; logistic regression and linear support vector machines by a similar wrapper around LIBLINEAR. In such cases, extending these methods with Python may not be possible [25].

Research results. Methods of analysis of text documents significantly expand the possibilities of cataloging text documents. This becomes possible due to the fact that these methods make it possible to record the kinship of texts, the meaningful proximity of words, which is not expressed simply in their similar spelling. Also, fewer and fewer processing tasks require pre-encoding arrays for model training. Another important advantage is that the application of these methods to document analysis makes it scalable: it makes no fundamental difference for a machine to analyze ten documents or one million, while with manual coding the complexity of the analysis largely depends on the size of the text library to be analyzed. It also

eliminates the sampling problem in document-based studies, as automated analysis allows for continuous sampling.

Also, along with the advantages of automated document analysis, there are also disadvantages. The main one is that the performance of each technical task (for example, classification) requires training of a new model sharpened for this task. Training such models requires arrays of texts labeled by researchers. This also creates a second drawback - the models are not 100% accurate and are highly dependent on the quality of the data and the algorithm they were trained with. The third disadvantage is the technical resources required for text processing.

Discussion of results. There are many programs that perform linguistic processing of text, but none of them use mechanisms for extracting content from textual information.

The main disadvantages of the considered methods and applications for text processing are:

- the cumbersomeness of the computational process, which is associated with the need for constant optimization of the input document;
- to bring the document to the form when it can be calculated, text simplifications are performed. Therefore, the probability that important words for the text will be simplified is quite high;
- there is a serious danger that too much attention to textual simplification may ignore the context of words, regardless of where they occur;
- when using some methods for large texts, the size of the final array may be too large, which will require a large number of parameters during its further processing;
- during training, the methods do not take into account the order of words, and also use a limited number of context words at each training iteration.

In the process of researching the methods of analysis of text documents, a general procedure of text analysis is obtained, an overview of existing approaches to solving the analysis task is given, the main approaches used in the task of text analysis are described, the stages of the analysis process are defined, and the most common mathematical methods of text document analysis are considered. The revealed features of use, advantages and disadvantages of the specified methods allow to draw a conclusion about the need for further improvement of analysis algorithms based on the specified methods, which would be simple to implement,

effective, have low computational costs during training and high quality of analysis in real tasks.

Based on the analysis of the obtained results, an algorithm will be developed, which, based on the methods analyzed above, should simplify the process of text analysis and allow the use of methods of computer analysis of text information to determine the attributes of a text document, based on which the analyzed text can be attributed to one or more groups, to which the specified attributes correspond.

Conclusions. The use of text information processing technologies is a promising area. The implementation of analytical processing of textual information is necessary in medicine, in business, and much more.

When processing textual information from a variety of disparate information resources, it is necessary to identify the following tasks:

- 1) selection of title, authors, keywords and construction of a conceptual model of the text;
- 2) integration into a full-text database;
- 3) search in full-text databases;
- 4) ensuring the relevance of the request;
- 5) reducing the amount of textual information and summarizing the text from several sources.

Search engines use an algorithm to search for duplicate text documents in order to index unique resources - that is, almost all branches of science and technology, due to the complexity of their organization, require intelligent processing of textual information. Only methods of processing textual information, the implementation of which involves all stages of analysis used in the analysis of language text, can claim to conduct a meaningful analysis of electronic text documents.

Almost all fields of knowledge, science and technology due to the complexity of their system organization, require intelligent data processing, therefore the processing of electronic documents today requires powerful mechanisms for analyzing textual information, which would ensure a complete and correct transition from natural language to mathematical models of text information processing.

As a result, it has been determined the need to develop a classification algorithm based on the mentioned approaches, such as lower case, root correction, stemming, lemmatization, stop word removal, normalization, which would be simple to implement, effective, have low computational costs during training and high

quality of classification in real tasks. Also, an approach to evaluating the thematic proximity of documents using feature space reduction is defined, and an algorithm for forming information-search attributes of documents for automatic clustering of documents is considered. The expediency of using methods of intelligent text analysis for this is considered.

In the future, an algorithm will be developed that will enable the application of methods of computer analysis of textual information, and software based on it. It will determine the attributes of a text document, based on which the analyzed text can be assigned to one or more groups, to which the specified attributes correspond.

References

- [1] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison Wesley, Reading, MA, 1998.
- [2] N. Kasyanchuk, and L. Tkachuk, "Protection of information in databases", in *Conf. VNTU of Electron. Sci. Publications, XLVIII Sci. and Tech. Conf. of the Faculty of Management and Information Security*, 2019, pp. 2419-2424 [in Ukrainian].
- [3] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*. 3rd ed. Morgan Kaufmann, 2011.
- [4] J. F. Luger, *Artificial Intelligence. Strategies and methods for solving complex problems*. 4th ed. Moscow: Izdat. Dom Williams, 2003.
- [5] T. Joachims, *Learning to Classify Text Using Support Vector Machines: Methods, Theory and Algorithms*. MA, USA: Kluwer Academic Publisher Norwel, 2002.
- [6] O. V. Havrylenko, Yu. O. Oliynyk, and G. V. Khanko, "Overview and analysis of text mining algorithms", *Project Management, System Analysis and Logistics*, no. 19, pp. 15-23, Kyiv, 2017 [in Ukrainian].
- [7] M. Lemke, and G. Wiedemann, *Text Mining in den Sozialwissenschaften*. Springer Fachmedien Wiesbaden, 2016, pp. 397-419.
- [8] I. V. Gushchin, and D. O. Sych, "Analysis of the influence of pre-processing of the text on the results of text classification", *Young Scientist*, no. 10, pp. 264-267, Kherson, 2018 [in Ukrainian].
- [9] G. Salton et al., "Automatic text structuring and summarization", *Information Processing & Management*, vol. 33, no. 2, pp. 193-207, 1997.
- [10] Z. Yao, Y. Sun, W. Ding, N. Rao, and H. Xiong, "Dynamic word embeddings for evolving semantic discovery", *WSDM 2018 Proc. 11th ACM Int. Conf. on Web Search and Data Mining*. Marina Del Rey, CA, USA, Febr. 5-9, 2018, pp. 673-681.
- [11] Word2Vec Implementation. [Online]. Available: <https://towardsdatascience.com/a-word2vec-implementation-using-numpy-and-python-d256cf0e5f28>.
- [12] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space", *arXiv:1301.3781*, 2013.
- [13] I. G. Oksanich, "Intellectual analysis of an array of text documents based on text mining technology", *Information Processing Systems*, pp. 139-143, Lutsk, 2013 [in Ukrainian].
- [14] A. Yu. Zubrytskyi, "Intellectual system of text research and analysis", M.S. thesis, National Technical University of Ukraine "Ihor Sikors'kyy Kyiv Polytechnic Institute, Kyiv, Ukraine, 2019 [in Ukrainian].
- [15] G. S. Linoff, and M. J. A. Berry, *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*, 3rd ed. NY, USA: Wiley Publishing inc., 2011.
- [16] S. Deerwester et al., *Indexing by Latent Semantic Analysis*. Chicago, IL, USA: Graduate Library School University of Chicago, 1990.
- [17] E. V. Bodyansky, and O. G. Rudenko, *Artificial Neural Networks: Architecture, Training, Application*. Kharkiv: TELETECH, 2004 [in Ukrainian].
- [18] D. W. Lande, *Search for Knowledge on the Internet. Professional Work*. NY, USA: Williams, 2005.
- [19] M. T. Hagan, H. B. Demuth, M. H. Beale, and O. De Jesús, *Neural Network Design*. 2014.
- [20] K. S. Jones, "A statistical interpretation of term specificity and its application in retrieval", *Journal of Documentation*, vol. 60, no. 5, pp. 493-502, MCB University Press, 2004.

- [21] A. Shalloway, and J. R. Trott, *Design Templates. A New Approach to Object-Oriented Analysis and Design*. NY, USA: Williams, 2002.
- [22] "Library of software components of text analysis technology". [Online]. Available: <https://www.analyst.ru/index.php?lang=rus&dir=content/downloads/>.
- [23] "Advego - content exchange №1". [Online]. Available: <https://advego.com/>.
- [24] DeepDive [Online]. Available: <http://deeplive.stanford.edu/>.
- [25] F. Pedregosa et al., "Scikit-learn: Machine learning in Python", *Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011.

Д. О. Якименко, аспірант (здобувач осв.-наук. ступеня доктора філософії),

Є. Ю. Катаєва, канд. техн. наук, доцент

Черкаський державний технологічний університет

б-р Шевченка, 460, м. Черкаси, 18006, Україна

МЕТОДИ ТА ЗАСОБИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ТЕКСТОВИХ ДОКУМЕНТІВ

В роботі проведено огляд методів аналізу та обробки електронних документів. Проаналізовано методи аналізу текстових документів для вирішення задачі визначення тематичної спорідненості текстів.

Виконано огляд існуючих підходів до вирішення задачі класифікації. Описано основні підходи, що використовуються в задачі класифікації текстів; визначено етапи процесу класифікації та розглянуто найпоширеніші методи класифікації текстових документів. Розглянуто основні підходи до попередньої обробки тексту: Нижній регістр, Коренева корекція, Стемінг, Лематизація, Видалення стоп-слова, Нормалізація. Розглянуто переваги та недоліки кожного підходу. Розглянуто процедуру зменшення розмірності набору ознак із поділом на підпроцеси: обирання ознак та виділення ознак. Розглянуто, в яких випадках кожен із підпроцесів є недоцільним для використання, та описано, які пошукові та фільтрові підходи і метрики є альтернативними або спорідненими для них.

Зроблено висновок щодо необхідності подальшого розроблення алгоритмів класифікації на базі зазначених методів, що були б простими в реалізації, ефективними, мали низькі обчислювальні витрати під час навчання та високу якість класифікації в реальних завданнях.

Визначено підхід до оцінки тематичної близькості документів з використанням редукції простору ознак і розглянуто алгоритм формування інформаційно-пошукових атрибутів документів для виконання автоматичної кластеризації документів. Розглянуто доцільність застосування для цього методів інтелектуального аналізу тексту.

Проаналізовано відкрите програмне забезпечення з використанням розглянутих методів.

Ключові слова: ключові слова, аналіз тексту, пошук, текстові документи, класифікація.

[0000-0002-2373-2429] **Є. П. Кириченко**

e-mail: kyrychenkojp@gmail.com

Черкаський інститут пожежної безпеки імені Героїв Чорнобиля
Національного університету цивільного захисту України
вул. Онопрієнка, 8, м. Черкаси, 18034, Україна

МЕТОДИКА ВИЗНАЧЕННЯ КРИТИЧНИХ ЗНАЧЕНЬ ПАРАМЕТРІВ ЗОВНІШНІХ ТЕРМІЧНИХ ДІЙ НА ПІРОТЕХНІЧНІ МЕТАЛООКСИДНІ ВИРОБИ В УМОВАХ ЕКСПЛУАТАЦІЇ

Практичне значення мають методики попередження вимушених передчасних пожежонебезпечних руйнувань виробів у випадку впливу зовнішніх термічних дій, які повинні базуватися на науково-обґрунтованих методах визначення критичних значень параметрів термічних впливів на вироби і технологічних параметрів зарядів сумішей, перевищення яких призводить до передчасних пожежонебезпечних руйнувань виробів. На базі отриманих експериментальних даних, математичних, експериментально-статистичних моделей, спеціалізованого програмного забезпечення у вигляді чотирьох стандартних пакетів прикладних програм розроблено науково-обґрунтовану методику визначення критичних діапазонів зміни параметрів зовнішніх термодій та керованих технологічних параметрів зарядів сумішей, перевищення яких призводить до пожежовибухонебезпечних руйнувань піротехнічних виробів. Практичне застосування розробленої методики для виробів на основі досліджуваних сумішей у вигляді засобів контролю і технологічних рекомендацій дозволяє знизити вірогідність їх руйнування в умовах експлуатації у 2...3 рази.

Ключові слова: пожежна безпека, піротехнічні суміші, процеси займання та горіння, порошки металевих палих та оксидів металів.

Вступ. Нині піротехнічні металооксидні вироби, основою яких є суміші металевих палих (магнію, алюмінію, титану, цирконію та ін.) та оксидів металів (оксидів міді, сурми, нікелю та ін.), широко використовуються як спалахувальні та запалювальні засоби для освітлювальних і сигнальних патронів та снарядів, ІЧ-засобів теплового захисту різних об'єктів, а також у вигляді засобів в артилерійських або бомбових дистанційних трубках та ін. Водночас з кожним роком в Україні та в усьому світі зростає кількість пожеж і вибухів у процесі обігу досліджуваних виробів, що призводить до руйнування об'єктів, травмування та загибелі людей, значних матеріальних збитків [9]. Велика кількість передчасних вибухонебезпечних спрацьовувань зазначених виробів мала катастрофічні наслідки (Мексика (2017 р.), Йорданія (2019 р.), Ліван (2020 р.) та ін.). Так, наприклад, у порту м. Бейрут (2020 р.) в результаті пожежі на складі піротехніки стався потужний вибух, що створив ударну хвилю, яка повністю зруйнувала припортову інфраструктуру, будівлі та споруди на відстані 10 км, загинуло сотні людей, тисячі – було поранено.

Зазначені випадки свідчать про те, що вироби спалахують і руйнуються з виникнен-

ням пожеж та вибухів внаслідок впливу зовнішніх термічних дій. Останні можуть бути обумовлені, наприклад, пожежею у складських приміщеннях, де зберігаються вироби, займанням близько розташованих об'єктів при транспортуванні, умовами пострілу та польоту виробів тощо.

Тому практичного значення набувають методики попередження руйнувань виробів з виникненням пожежі при впливі зовнішніх термічних дій. Основою цих методик є методи визначення значень параметрів термічних впливів і технологічних характеристик зарядів сумішей, перевищення яких призводить до пожежонебезпечних руйнувань виробів.

На цей момент на основі проведених достатньо широких теоретико-експериментальних досліджень зазначені методи розроблено тільки для широко використовуваних у піротехніці сумішей металевих палих і нітратовмісних окиснювачів (нітратів лужних та лужноземельних металів) [5-8]. Розроблено методи для визначення критичних значень параметрів зовнішніх термічних впливів (теплових потоків, часів їх дії, швидкостей обдуву потоком повітря, режимів обтікання та ін.) [10-13]. Запропоновано метод для визначення критичних діапазонів зміни швидкості горін-

ня зазначених сумішей в умовах підвищених температур нагріву (до 600...700 K) і зовнішніх тисків (до 10^7 Па) [14-16]. Сформульовано практичні рекомендації з використання розроблених методів у піротехнічному виробництві виробів на основі зазначених сумішей та при їх експлуатації, що дозволяють знижувати вірогідність пожежонебезпечних руйнувань виробів при наявності екстремальних термічних впливів [16-18]. Щодо досліджуваних піротехнічних металооксидних виробів, то нині зазначені методи для них відсутні.

Мета та задачі досліджень. *Мета роботи:* розробка методики визначення значень характеристик термічних впливів та керованих технологічних параметрів зарядів сумішей, що призводять до передчасного займання та вибухового розвитку горіння піротехнічних сумішей з порошків магнію, алюмінію та оксидів металів, та, у кінцевому підсумку, до пожежонебезпечного руйнування виробів.

Задачами досліджень є:

- сформувати програмний комплекс, що дозволяє вирішувати типові інженерно-технологічні задачі пожежної безпеки різних піротехнічних сумішей;
- розробити алгоритм практичного використання методики;
- оцінити підвищення пожежної безпеки при використанні методики.

Виклад основного матеріалу. В основу методики покладено розроблені раніше [18, 19] математичні та експериментально-статистичні моделі, а також стандартне спеціалізоване програмне забезпечення у вигляді чотирьох пакетів прикладних програм ПАКЕТ1 – ПАКЕТ4 за сучасними чисельними методами, використовуючи які, формується база даних по критичних значеннях, перевищення яких призводить до передчасного розвитку вибухового горіння, величин теплових потоків та часів їх дії, а також технологічних параметрах зарядів (співвідношеннях компонентів та їх дисперсності, коефіцієнтах ущільнення сумішей, геометричних розмірах виробів).

Комплекс моделей. На рисунку 1 зображено комплекс розроблених математичних та експериментально-статистичних моделей [18, 19], розрахунки за якими дозволяють сформулювати базу даних, на основі якої й було розроблено зазначену вище методику.

Спеціалізований програмний комплекс. Для проведення розрахунків за моделями використовувався спеціалізований програмний комплекс СПК, що містить стандартні пакети

прикладних програм (ППП) ПАКЕТ1 – ПАКЕТ4 з основних методів розв’язання обчислювальних задач (рисунк 2).

Програмний комплекс складається з чотирьох спеціалізованих PPP, що містять 80 програмних модулів з сучасних чисельних методів розв’язання лінійних і нелінійних, одиничних і систем алгебраїчних та трансцендентних рівнянь, одиничних і систем звичайних диференціальних рівнянь 1-го порядку і 2-го порядку, диференціальних рівнянь у часткових похідних 2-го порядку, крайових задач для звичайних і диференціальних рівнянь у часткових похідних 2-го порядку, інтегральних рівнянь Вольтера II роду, а також методів чисельного інтегрування функцій, регресійного аналізу, інтерполяції і статистичної обробки експериментальних даних. Ефективність зазначених PPP було встановлено в результаті проведення широких чисельних експериментів (у діалоговому режимі було вирішено більше 300 інженерно-технологічних задач, властивих процесам горіння піротехнічних сумішей).

Алгоритм практичного використання методики. Згідно з розробленим алгоритмом (рисунк 3) на першому етапі створюється вихідна база даних, яка містить параметри зовнішніх термодій (густина теплового потоку q_n , час його дії t , швидкість обдуву потоком повітря V , критерій Рейнольдса Re , температура нагріву T_0 та тиск P , при яких починається горіння зарядів сумішей), технологічні характеристики зарядів сумішей (природа окиснювача та металевго пального, відносний вміст металевго пального ζ_m у суміші, середній розмір частинок порошку металевго пального d_m та порошку окиснювача $d_{ок}$), механічні та теплофізичні параметри зарядів сумішей (геометрична форма та розміри, густина, теплоємність та коефіцієнт теплопровідності), отриманий комплекс математичних та експериментально-статистичних моделей, пакети прикладних програм ПАКЕТ1 – ПАКЕТ4, що входять у СПК, з метою проведення розрахунків на ПК. На другому етапі вводяться значення T_n (температура поверхні заряду суміші) та T_3 (температура займання металевго пального у продуктах розкладання окиснювача), при перевищенні яких ($T_n > T_3$), що відповідає критичним зовнішнім термічним впливам ($q_n > q_n^*$, $t > t^*$, $V > V^*$, $Re > Re^*$), відбувається прискорення саморозігрівання сумішей, передчасне їх займання та розвиток вибухового горіння. Потім вибирається мо-

дель зовнішніх термічних впливів на заряд суміші, використовуючи ППП, розраховуються значення параметра T_n , далі вони порівнюються з їх граничними значеннями і знаходяться критичні значення q_n^* , t^* , V^* та Re^* , при яких вироби руйнуються. Проводяться вимірювання температури поверхні T_n , теплового потоку q_n , швидкості обдуву V та критерію Рейнольдса Re в окремих точках заряду. Результати вимірювань зіставляються з результатами розрахунків та визначаються помилки розрахункових методів знаходження критичних параметрів q_n^* , t^* , V^* та Re^* . Якщо виявлено істотні розбіжності (більше 10...12 %), здійснюється коригування моделей. На третьому етапі, якщо виріб залишається у працездатному стані у жорстких умовах експлуатації, тобто не відбулись перегрів корпусу та передчасне спрацювання заряду суміші, відсутні локальне або повне руйнування, вводяться значення температури займання T_z , часу згоряння частинок металевго пального τ_z , швидкості горіння сумішей u та їх граничні значення (τ_z^* , T_z^* , $u_{(\zeta)}^*$) (швидкість горіння на верхній концентраційній межі, перевищення якої веде до вибухового розвитку процесу

горіння), $u_{(\zeta)}^*$ (швидкість горіння на нижній концентраційній межі, перевищення якої призводить до затухання процесу горіння). Далі вибирається математична або експериментально-статистична модель та, використовуючи ППП, проводиться розрахунок τ_z , T_z , u при заданих значеннях керованих параметрів (ζ_m , d_m , $d_{ок}$, T_0 , P), після чого значення параметрів зіставляються з їх граничними значеннями (перевірка виконання умов: $\tau_z < \tau_z^*$, $T_z < T_z^*$, $u < u_{(\zeta)}^*$, $u > u_{(\zeta)}^*$). Якщо зазначені умови не виконуються, проводиться коригування значень керованих параметрів і визначаються критичні діапазони їх зміни: $\zeta_m < (\zeta_m)_{ВМГ}$, $\zeta_m > (\zeta_m)_{НМГ}$, $d_m < d_m^*$ (або $d_m > d_m^*$), $d_{ок} < d_{ок}^*$ (або $d_{ок} > d_{ок}^*$), $T_0 > T_0^*$ (або $T_0 < T_0^*$), $P > P^*$ (або $P < P^*$). Наступними кроками є проведення вибірових вимірювань температури займання та часу горіння, швидкості горіння, зіставлення з їх розрахунковими значеннями, і визначення похибки розрахункових методів знаходження критичних діапазонів зміни керованих параметрів. Коригування моделей проводиться у разі виявлення значних розбіжностей між розрахунковими та експериментальними даними (більше 10...12 %).



Рисунок 1 – Комплекс розроблених математичних та експериментально-статистичних моделей для формування бази даних



На завершення слід відзначити, що розроблена науково обґрунтована методика дає можливість при розробці піротехнічних ви-

бів, основою яких є суміші з порошків магнію, алюмінію та оксидів металів, формувати в автоматизованому режимі базу даних по критичних значеннях параметрів термічних дій, перевищення яких призводить до передчасного займання зарядів сумішей, прискореного розвитку їх горіння та зрештою до вибухонебезпечного руйнування виробів.

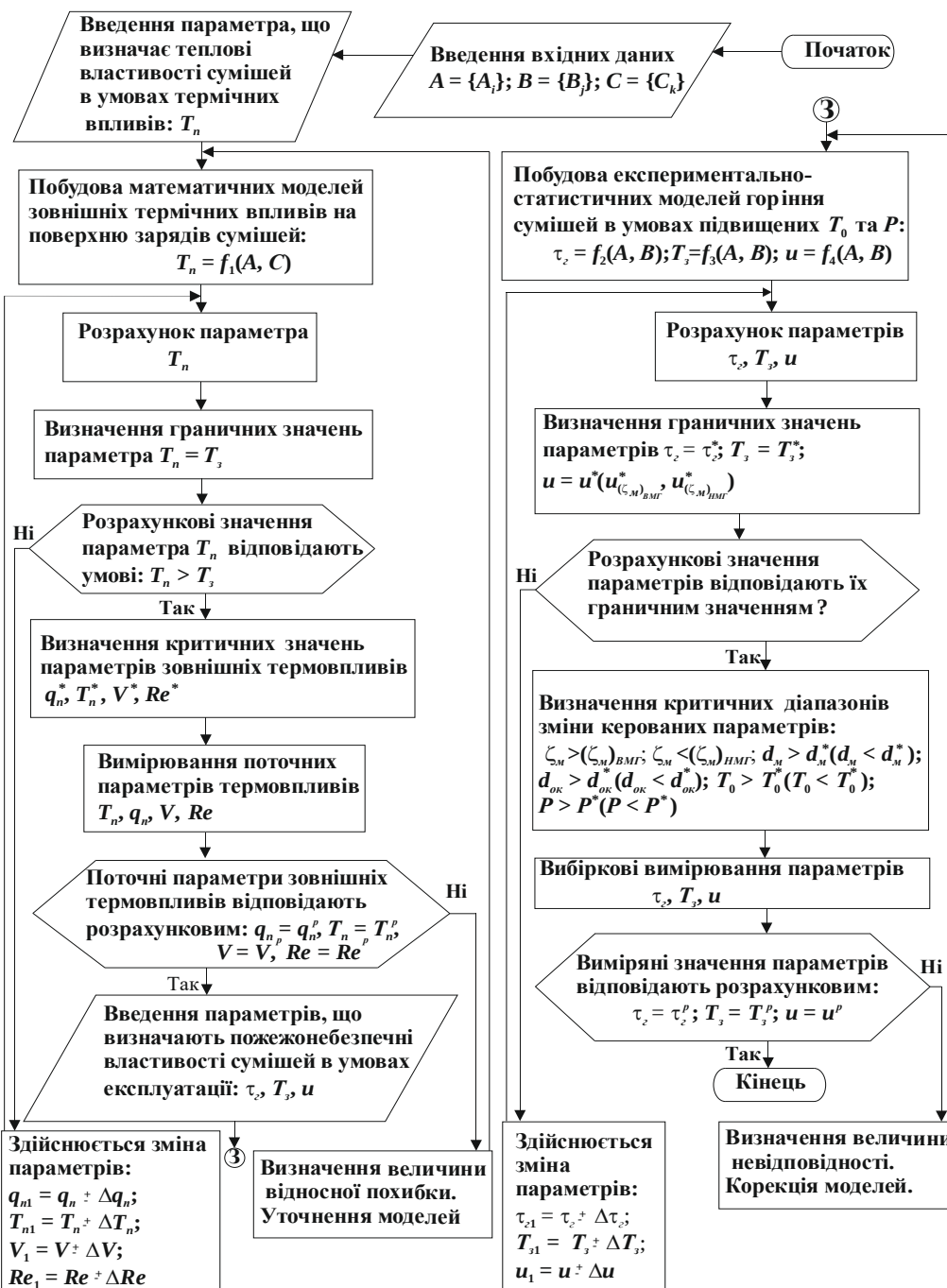


Рисунок 3 – Алгоритм визначення критичних значень параметрів зовнішніх термічних дій (теплових потоків та часів їх дії, швидкостей обдуву потоком повітря та критеріїв Рейнольдса) на піротехнічні вироби з зарядами сумішей, що дозволяє попереджати можливі передчасні їх спрацювання та пожежонебезпечні руйнування у процесі експлуатації: A – параметри зовнішніх термічних дій; B – керовані технологічні параметри сумішей; C – механічні, теплофізичні та фізико-хімічні характеристики зарядів сумішей

Результати досліджень. Для запобігання займання та подальшого пожежонебезпечного руйнування піротехнічних виробів з зарядами досліджуваних сумішей в умовах зовнішнього термічного впливу необхідним є запобігання утворенню у горючому середовищі осередків екзотермічного окиснення частинок магнію та алюмінію в продуктах розкладання оксидів металів. Останнє забезпечується зниженням горючості сумішей, що досягається шляхом варіювання їх технологічних параметрів (співвідношення та дисперсності компонентів, коефіцієнта ущільнення, розмірів заряду) [20].

Для оцінки рівня пожежної безпеки виробів в умовах зовнішньої термічної дії (див. ДСТУ 8829:2019 п. 3. Розрахунок вірогідності появи джерела запалювання (ініціювання горіння та вибуху [4])) розраховується вірогідність W появи джерела запалювання у піротехнічному виробі, яке призводить до передчасного горіння заряду суміші виробу. Враховується, що при експлуатації піротехнічних виробів неможливо передбачити умови виникнення зовнішніх термічних дій на їх поверхню, які можуть призвести до появи у заряді сумішей осередків екзотермічного окиснення частинок металу в продуктах розкладання оксидів металів. Тому згідно з ДСТУ 8829:2019 [4] для розрахунку W використовується вираз

$$W = 1 - e^{-t/\tau}, \quad (1)$$

де t – час зберігання (транспортування) піротехнічного виробу, год; $\tau = 3,03 \cdot 10^4 \cdot E_0^{1,2}$ – середній час зберігання (транспортування) до появи джерела запалювання, год; E_0 – мінімальна енергія запалювання суміші (утворення осередків екзотермічного окиснення), Дж. Величину E_0 можна визначити як мінімальну енергію, необхідну для нагріву об'єму реакційної зони к-фази заряду від початкової температури (T_0 , К) до температури займання частинок металевого пального у продуктах розкладання оксидів металів (T_z , К) в умовах зовнішнього термовпливу [18]:

$$E_0 = \rho_c \cdot c_c \cdot V_c \cdot (T_z - T_0), \quad (2)$$

де ρ_c , c_c – густина (кг/м³) та питома теплоємність (Дж/кг·К) суміші; V_c – об'єм реакційної зони к-фази суміші, в якій відбувається термічне розкладання оксиду металу, окиснення та спалахування частинок металевого пального, м³.

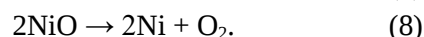
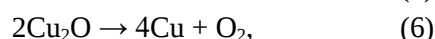
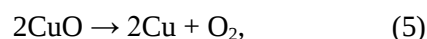
Оскільки заряд має діаметр d та висоту h , а товщина реакційної зони к-фази становить δ , у першому наближенні можна вважати [19]

$$V_c \approx 0,5 \pi d \cdot h \cdot \delta. \quad (3)$$

Підставляючи (3) у (2), отримуємо

$$E_0 \approx \rho_c \cdot c_c \cdot 0,5 \pi d \cdot h \cdot \delta \cdot (T_z - T_0). \quad (4)$$

Для знаходження T_z скористаємося результатами, отриманими у роботах [19, 20]. Для використання цих даних враховано, що при температурах T_p , які відповідають реакційній зоні к-фази ($T_p = 1400 \dots 1800$ К) досліджувані окиснювачі розкладаються згідно з реакціями [18, 19]:



Молекулярні маси хімічних елементів, що входять у (5)–(8): $\mu_{\text{CuO}} = 79,54$, $\mu_{\text{Cu}} = 63,54$, $\mu_{\text{O}_2} = 32$, $\mu_{\text{Cu}_2\text{O}} = 143,08$, $\mu_{\text{Sb}_2\text{O}_3} = 291,52$, $\mu_{\text{Sb}} = 121,76$, $\mu_{\text{NiO}} = 74,71$, $\mu_{\text{Ni}} = 58,71$. Тоді відносні масові вмісти O_2 у продуктах розкладання окиснювачів: $C_{\text{O}_2} \approx 0,11$ (окиснювач Sb_2O_3); $C_{\text{O}_2} \approx 0,14$ (окиснювач Cu_2O); $C_{\text{O}_2} \approx 0,21$ (окиснювач NiO); $C_{\text{O}_2} \approx 0,25$ (окиснювач CuO). Враховуючи також, що розглядається початковий період займання частинок магнію та алюмінію (зовнішній тиск $P = 10^5$ Па), отримуємо, що на величину T_z найсильніше впливає дисперсність порошку металевого пального (рисунки 4, 5). Відносна похибка для кривих, що апроксимують експериментальні дані, становить 5...7 %.

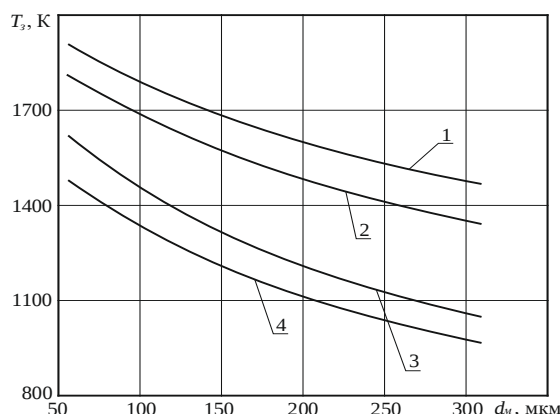


Рисунок 4 – Залежності температури T_z від середнього розміру частинок порошку магнію та природи окиснювача ($P = 10^5$ Па): 1 – суміш $\text{Mg} + \text{Sb}_2\text{O}_3$; 2 – суміш $\text{Mg} + \text{Cu}_2\text{O}$; 3 – суміш $\text{Mg} + \text{NiO}$; 4 – суміш $\text{Mg} + \text{CuO}$

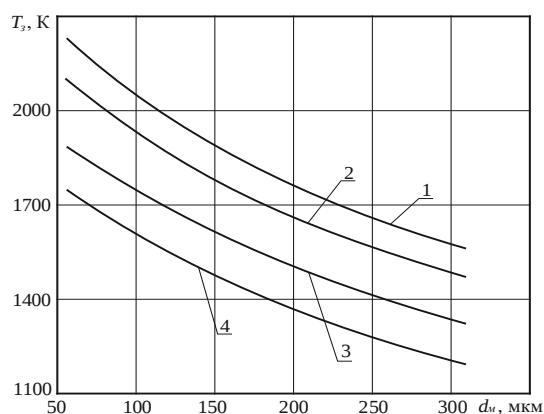


Рисунок 5 – Залежності температури T_z від середнього розміру частинок порошку алюмінію та природи окиснювача ($P = 10^5$ Па): 1 – суміш $\text{Al} + \text{Sb}_2\text{O}_3$; 2 – суміш $\text{Al} + \text{Cu}_2\text{O}$; 3 – суміш $\text{Al} + \text{NiO}$; 4 – суміш $\text{Al} + \text{CuO}$

Обговорення результатів. З результатів, зображених на рисунках 4, 5, випливає, що для досліджуваних діапазонів зміни середніх розмірів частинок порошку металевого пального ($d_m = 56...305$ мкм – для магнію та $d_m = 4...310$ мкм – для алюмінію) збільшення d_m призводить до зменшення T_z : для сумішей магній + оксиди металів – у 1,3...1,5 разу; для сумішей алюміній + оксиди металів – у 1,4...1,6 разу. Це пояснюється тим, що зі збільшенням d_m значно зростає площа поверхні ($\sim d_m^2$), на якій відбуваються гетерогенні екзотермічні реакції окиснення, що призводить до займання частинок при більш низьких T_z внаслідок швидкого їх нагріву. Заміна кисневмісного окиснювача на окиснювач, при термічному розкладанні якого вміст кисню зменшується (наприклад, при заміні CuO на Cu_2O значення C_{O_2} зменшуються у 2,3 разу), призводить до збільшення T_z у 1,3...1,4 разу (див.

рисунки 4, 5). Таким чином, використання дрібнодисперсних порошоків магнію та алюмінію, а також оксидів металів зі зменшеним вмістом кисню у їх продуктах розкладання на стадії виготовлення зарядів досліджуваних сумішей дозволить зменшити передчасне ініціювання зарядів сумішей в умовах зовнішнього термовпливу, тобто підвищити пожежну безпеку піротехнічних виробів.

Для кількісної оцінки сказаного вище розрахуємо вірогідність W ініціювання заряду досліджуваної суміші при термічному впливі, тобто вірогідність пожежонебезпечного руйнування піротехнічних виробів.

Розраховуючи за формулами (1)–(4) та використовуючи дані, зображені на рисунках 4, 5 на прикладі чотирьох стехіометричних зарядів сумішей (таблиця 1), з урахуванням відомих даних ($T_0 = 293$ К; $P = 10^5$ Па; $h = 3 \cdot 10^{-2}$ м; $d = 2 \cdot 10^{-2}$ м; $\delta = 10^{-5}...10^{-6}$ м; $t = 8,76 \cdot 10^3$ год [18]), отримуємо, що вірогідність W зі збільшенням d_m суттєво збільшується (рисунки 6, 7): для виробу 1 – у 4,1 разу; для виробу 2 – у 3,5 разу; для виробу 3 – у 2,8 разу; для виробу 4 – у 2,2 разу.

Таким чином, шляхом зміни дисперсності порошоків металевого пального, що використовуються в зарядах досліджуваних сумішей, можна підвищувати пожежну безпеку піротехнічних виробів на їх основі в умовах зовнішніх термічних впливів більше, ніж у 2...3 рази.

Крім цього, замінюючи один оксид металу на інший, що має менший вміст вільного кисню (наприклад, заміна CuO на Cu_2O), також можна підвищити пожежну безпеку піротехнічних виробів у 1,5...1,8 разу.

Таблиця 1 – Перелік серійних піротехнічних виробів з досліджуваними зарядами сумішей, для яких розраховувалась вірогідність W згідно з ДСТУ 8829:2019 [4]

№ з/п	Призначення	Склад заряду суміші	Відносний масовий вміст компонентів у суміші, %
1	Піротехнічні спалахувачі для освітлювальних патронів та снарядів	Mg	21
		CuO	79
2	Піротехнічні спалахувачі для сигнальних засобів	Mg	31
		Cu ₂ O	69
3	Піротехнічні спалахувачі для трасерів білого вогню	Al	27
		CuO	73
4	Піротехнічні запалювальні вироби	Al	75
		Cu ₂ O	25

Примітка. Технологічні параметри зарядів сумішей: $K_y = 0,95...0,96$; $d_m = 54...310$ мкм; $d_{ок} = 30...50$ мкм; $h = 3 \cdot 10^{-2}...4 \cdot 10^{-2}$ м; $d = 1,5 \cdot 10^{-2}...2 \cdot 10^{-2}$ м

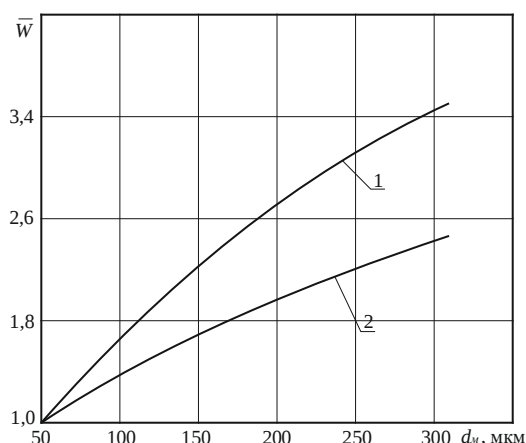


Рисунок 6 – Залежність відносної вірогідності $\bar{W}$ ($\bar{W} = \frac{W}{W_{Mg}^*}$, де W_{Mg}^* – значення вірогідності для мінімального розміру частинок порошків магнію з використовуваних на практиці діапазонів його зміни [18, 19]) пожежонебезпечного руйнування піротехнічних виробів на основі сумішей магній + оксиди металів від дисперсності порошку металевого пального: 1 – вироби на основі сумішей Mg + CuO; 2 – вироби на основі сумішей Mg + Cu₂O (див. таблицю 1)

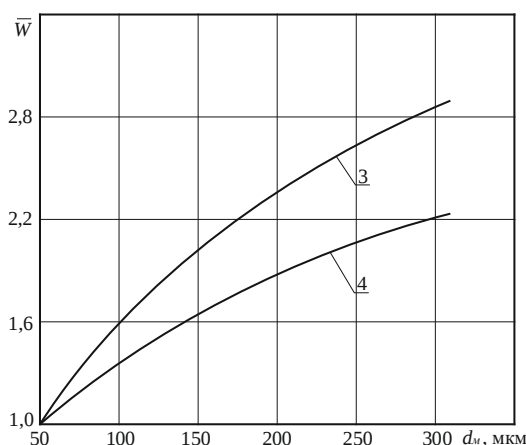


Рисунок 7 – Залежність відносної вірогідності $\bar{W}$ ($\bar{W} = \frac{W}{W_{Al}^*}$, де W_{Al}^* – значення вірогідності для мінімального розміру частинок порошків алюмінію з використовуваних на практиці діапазонів його зміни [18, 19]) пожежонебезпечного руйнування піротехнічних виробів на основі сумішей алюміній + оксиди металів від дисперсності порошку металевого пального: 3 – вироби на основі сумішей Al + CuO; 4 – вироби на основі сумішей Al + Cu₂O (див. таблицю 1)

Висновки:

1. Сформовано спеціалізований програмний комплекс, що дозволяє в автоматизованому режимі вирішувати типові інженерно-технологічні задачі пожежної безпеки різних піротехнічних сумішей.

2. Розроблено нову науково обґрунтовану методику, що базується на отриманих математичних, експериментально-статистичних моделях та спеціалізованому програмному комплексі, яка дає можливість створювати базу даних по критичних діапазонах зміни параметрів зовнішніх термічних впливів та керованих технологічних параметрах зарядів сумішей, перевищення яких призводить до пожежонебезпечних руйнувань виробів.

3. Показано, що практичне використання розробленої методики у вигляді засобів контролю та технологічних рекомендацій дозволяє знижувати пожежну небезпеку піротехнічних виробів (освітлювальних, сигнальних, трасувальних патронів та снарядів, запалювальних засобів тощо) в умовах зовнішніх термічних дій у 2...3 рази.

Перспективи подальших досліджень.

Перспективним напрямом є подальше практичне застосування отриманих результатів, що полягає у розробці та впровадженні:

- науково обґрунтованої методики визначення критичних значень параметрів зовнішніх термічних впливів і технологічних характеристик зарядів піротехнічних сумішей з порошків магнію, алюмінію та оксидів металів, перевищення яких призводить до їх передчасного спрацьовування та пожежонебезпечного руйнування виробів;

- засобів контролю і технологічних рекомендацій, що дозволяють знизити вірогідність пожежонебезпечних руйнувань піротехнічних виробів при зберіганні та транспортуванні з урахуванням зовнішніх термічних впливів у 2...3 рази.

Список використаних джерел

- [1] V. V. Kovalyshyn, V. M. Marych, Y. M. Novitskyi, B. M. Gusar, V. V. Chemetskiy, and O. L. Mirus, "Improvement of a discharge nozzle damping attachment to suppress fires of class D", *Eastern-European Journal of Enterprise Technologies*, no. 5 (95), pp. 68-76, 2018.

- [2] В. М. Фатєєв, Ю. П. Приходько, та Л. І. Таборов, *Піротехніка*. Київ, Україна: Наукова думка, 2017.
- [3] ДСТУ 8828:2019. Пожежна безпека. Загальні положення. Чинний від 01.01.2020. Київ: ДП УкрНДНЦ, 2019.
- [4] ДСТУ 8829:2019. Пожежовибухонебезпечність речовин і матеріалів. Номенклатура показників і методи їхнього визначення. Класифікація. Чинний від 01.01.2020. Київ: ДП УкрНДНЦ, 2019.
- [5] В. М. Баланюк, Н. М. Козяр, та О. І. Гарасим'юк, "Застосування газоаерозольно-порошкових вогнегасних сумішей для захисту від запалювальних сумішей", *Science Rise*, № 2 (22), с. 11-14, 2016.
- [6] В. М. Марич, В. В. Ковалишин, та Я. Б. Кирилів, "Оптимізація складу вогнегасних порошків для гасіння магнію", на *Міжнар. наук.-практ. конф. Теорія і практика гасіння пожеж та ліквідації надзвичайних ситуацій*. Черкаси, 2018.
- [7] В. В. Ковалишин, В. М. Марич, Т. М. Войтович, та Б. М. Гусар, "Використання екологічно прийнятних вогнегасних речовин", на *Міжнар. наук.-практ. конф. Екологічна безпека як основа сталого розвитку суспільства. Європейський досвід і перспективи*. ДУБЖД, 2018.
- [8] В. М. Марич, В. В. Ковалишин, та Я. Б. Кирилів, "Дослідження хімічних речовин як складників вогнегасних порошків для гасіння магнію та його сплавів", на *Міжнар. наук.-практ. конф. Теорія і практика гасіння пожеж та ліквідації надзвичайних ситуацій*. Черкаси, 2017.
- [9] Дослідження тенденцій і закономірностей динаміки основних показників статистики пожеж в Україні за територіальним принципом: звіт про наук.-дослідну роботу. Київ: УкрНДНЦ, 2018.
- [10] O. Dibrova, O. Kyrychenko, R. Motrychuk, M. Tomenko, and V. Melnyk, "Fire safety improvement of pyrotechnic nitrate-metal mixtures under external thermal conditions", *Technology audit and production reserves*, no. 1/1 (51), pp. 44-49, 2020.
- [11] О. В. Кириченко, О. С. Діброва, Р. Б. Мотрічук, В. А. Ващенко, С. О. Колінко, Т. І. Бутенко, та В. В. Цибулін, "Визначення критичних режимів розвитку процесів горіння піротехнічних нітратно-металевих сумішей в умовах зовнішніх термічних дій", *Вісник Черкаського державного технологічного університету*, № 2, с. 123-133, 2020.
- [12] O. V. Kyrychenko, O. S. Dibrova, R. B. Motrichuk, V. A. Vashchenko, and S. O. Kolinko, "Investigation of ignition and combustion of particles of aluminum and magnesium alloys in the decomposition products of solid pyrotechnic fuels", *Scientific Bulletin Civil Protection and Fire Safety*, no. 2 (8), pp. 81-85, 2019.
- [13] О. В. Кириченко, О. С. Діброва, Р. Б. Мотрічук, С. О. Тищенко, та В. В. Цибулін, "Визначення допустимих режимів нагріву піротехнічних сумішей при їх експлуатації", *Вісник Черкаського державного технологічного університету*, № 2, с. 5-11, 2018.
- [14] О. С. Діброва, О. В. Кириченко, Р. Б. Мотрічук, та В. А. Ващенко, "Підвищення пожежної безпеки піротехнічних нітратно-металевих сумішей в умовах зовнішніх термічних дій", *Internauka*, № 5/5799, 2020. [Електронний ресурс]. Режим доступу: <http://www.inter-nauka.com>.
- [15] О. С. Діброва, О. В. Кириченко, Р. Б. Мотрічук, та В. А. Ващенко, "Закономірності впливу технологічних параметрів на пожежну безпеку піротехнічних нітратно-титанових сумішей в умовах зовнішніх термічних дій", *Internauka*, № 5/5798, 2020. [Електронний ресурс]. Режим доступу: <http://www.inter-nauka.com>.
- [16] О. В. Кириченко, О. С. Діброва, та Р. Б. Мотрічук, "Вплив технологічних параметрів на залежності швидкості розвитку процесу горіння піротехнічних сумішей", на *XI Міжнар. наук.-практ. конф. "Теорія і практика гасіння пожеж та ліквідації надзвичайних ситуацій"*. Черкаси: Черкас. ін-т пожежної безпеки ім. Героїв Чорнобиля, 2020.
- [17] Р. Б. Мотрічук, О. В. Кириченко, В. А. Ващенко, С. О. Колінко, Т. І. Бутенко, Є. П. Кириченко, та В. В. Цибулін, "Закономірності впливу технологічних параметрів та зовнішніх чинників на температуру та склад продуктів згорання піротехнічних нітратно-металевих сумішей", *Вісник Черкаського державного технологічного університету*, № 4, с. 131-142, 2020.
- [18] Є. П. Кириченко, В. М. Гвоздь, В. А. Ващенко, О. В. Кириченко, О. О. Дядюшенко, та В. П. Мельник, "Закономірності

- впливу технологічних параметрів і зовнішніх чинників на температуру займання та час згоряння частинок магнію та алюмінію в продуктах розкладання оксидів металів", *Цивільний захист та пожежна безпека*, № 2 (12), с. 112-122, 2021.
- [19] Є. П. Кириченко, "Дослідження процесів зовнішніх термоударних дій на піротехнічні металооксидні вироби в умовах пострілу та польоту", *Пожежна безпека: теорія і практика*: зб. наук. пр. ЧПБ ім. Героїв Чорнобиля, т. 5, № 2, с. 87-102, 2021.
- [20] Є. П. Кириченко, В. В. Ковалишин, В. М. Гвоздь, В. А. Ващенко, С. О. Колінько, та В. В. Цибулін, "Дослідження механізму та розробка моделі розвитку процесу горіння піротехнічних сумішей металеве пальне+оксид металу при зовнішніх термічних діях", *Вісник Черкаського державного технологічного університету*, № 4, с. 68-82, 2021.
- ### References
- [1] V. V. Kovalyshyn, V. M. Marych, Y. M. Novitskiy, B. M. Gusar, V. V. Chemetskiy, and O. L. Mirus, "Improvement of a discharge nozzle damping attachment to suppress fires of class D", *Eastern-European Journal of Enterprise Technologies*, no. 5 (95), pp. 68-76, 2018.
- [2] V. M. Fateev, Y. P. Prikhodko, and L. I. Taborov, *Pyrotechnics*. Kyiv, Ukraine: Naukova dumka, 2017 [in Ukrainian].
- [3] DSTU 8828:2019. Fire Security. Terms. Effective from 2020, Jan. 01. Kyiv: UkrNDNC, 2019 [in Ukrainian].
- [4] DSTU 8829:2019. Fire and explosion hazard of substances and materials. Nomenclature of indicators and methods of their definition. Classification. Effective from 2020, Jan. 01. Kyiv: UkrNDNC, 2019 [in Ukrainian].
- [5] V. M. Balanyuk, N. M. Kozyar, and O. I. Garasimyuk, "Application of aerosol-powder fire-extinguishing mixtures for protection against incendiary mixtures", *Science Rise*, no. 2 (22), pp. 11-14, 2016 [in Ukrainian].
- [6] V. M. Marich, V. V. Kovalishin, and Ya. B. Kirillov, "Optimization of fire extinguishing powders for magnesium extinguishing", in *Proc. Int. Sci.-Pract. Conf. Theory and practice of firefighting and emergency response*. Cherkasy, 2018 [in Ukrainian].
- [7] V. V. Kovalishyn, V. M. Marich, T. M. Voityvych, and B. M. Husar, "Use of environmentally friendly fire extinguishers", in *Proc. Int. Sci.-Pract. Conf. Environmental security as a basis for sustainable development of society. European experience and prospects*. DUBZD, 2018 [in Ukrainian].
- [8] V. M. Marich, V. V. Kovalishyn, and Ya. B. Kirillov, "Investigation of chemicals as components of fire-extinguishing powders for extinguishing magnesium and its alloys", in *Proc. Int. Sci.-Pract. Conf. Theory and practice of firefighting and emergency response*. Cherkasy, 2017 [in Ukrainian].
- [9] Study of trends and patterns of dynamics of the main indicators of fire statistics in Ukraine on a territorial basis: Report on research work. Kyiv: UkrNDICZ, 2018 [in Ukrainian].
- [10] O. Dibrova, O. Kyrychenko, R. Motrychuk, M. Tomenko, and V. Melnyk, "Fire safety improvement of pyrotechnic nitrate-metal mixtures under external thermal conditions", *Technology audit and production reserves*, no. 1/1 (51), pp. 44-49, 2020.
- [11] O. V. Kirichenko, O. S. Dibrova, R. B. Motrichuk, V. A. Vashchenko, S. O. Kolinko, T. I. Butenko, and V. V. Tsybulin, "Determination of critical modes of development of combustion processes of pyrotechnic nitrate-metal mixtures in the conditions of external thermal actions", *Vіsnyk Cherkaskogo derzhavnogo tekhnologіchnogo unіversitetu*, no. 2, pp. 123-133, 2020 [in Ukrainian].
- [12] O. V. Kyrychenko, O. S. Dibrova, R. B. Motrichuk, V. A. Vashchenko, and S. O. Kolinko, "Investigation of ignition and combustion of particles of aluminum and magnesium alloys in the decomposition products of solid pyrotechnic fuels", *Scientific Bulletin Civil Protection and Fire Safety*, no. 2 (8), pp. 81-85, 2019.
- [13] O. V. Kirichenko, O. S. Dibrova, R. B. Motrichuk, E. O. Tishchenko, and V. V. Tsybulin, "Determination of permissible modes of heating of pyrotechnic mixtures during their operation", *Vіsnyk Cherkaskogo derzhavnogo tekhnologіchnogo unіversitetu*, no. 2, pp. 5-11, 2018 [in Ukrainian].
- [14] O. S. Dibrova, O. V. Kirichenko, R. B. Motrichuk, and V. A. Vashchenko, "Improving

- fire safety of pyrotechnic nitrate-metal mixtures in the conditions of external thermal actions", *Intenauka*, no. 5/5799, 2020. [Online]. Available: <http://www.intenauka.com> [in Ukrainian].
- [15] O. S. Dibrova, O. V. Kirichenko, R. B. Motrichuk, and V. A. Vashchenko, "Regularities of influence of technological parameters on fire safety of pyrotechnic nitrate-titanium mixtures in the conditions of external thermal actions", *Intenauka*, no. 5/5798, 2020. [Online]. Available: <http://www.intenauka.com> [in Ukrainian].
- [16] O. V. Kirichenko, O. S. Dibrova, and R. B. Motrichuk, "Influence of technological parameters on the dependences of the rate of development of the combustion process of pyrotechnic mixtures", in *Proc. XI Int. Sci.-Pract. Conf. Theory and practice of firefighting and emergency response*. Cherkasy: Cherkas. in-t pozhezhnoyi bezpeky im. Heroyiv Chornobylya, 2020 [in Ukrainian].
- [17] R. B. Motrichuk, O. V. Kirichenko, V. A. Vashchenko, S. O. Kolinko, T. I. Butenko, E. P. Kirichenko, and V. V. Tsybulin, "Regularities of influence of technological parameters and external factors on temperature and composition of combustion products of pyrotechnic nitrate-metal mixtures", *Vísnyk Cherkaskogo derzhavnogo tekhnologíchnogo uníversitetu*, no. 4, pp. 131-142, 2020 [in Ukrainian].
- [18] E. P. Kirichenko, V. M. Gvozdz, V. A. Vashchenko, O. V. Kirichenko, O. O. Dyadyushenko, and V. P. Melnik, "Regularities of the influence of technological parameters and external factors on the ignition temperature and combustion time of magnesium and aluminum particles in the decomposition products of metal oxides", *Tsyvilnyy zakhyst ta pozhezhna bezpeka*, no. 2 (12), pp. 112-122, 2021 [in Ukrainian].
- [19] E. P. Kirichenko, "Research of processes of external thermal shock actions on pyrotechnic metal oxide products in the conditions of a shot and flight", *Pozhezhna bezpeka: teoriya i praktyka*: coll. of sci. papers of Cherkasy Institute of Fire Safety named after Chornobyl Heroes, vol. 5, no. 2, pp. 87-102, 2021 [in Ukrainian].
- [20] E. P. Kirichenko, V. V. Kovalishin, V. M. Gvozdz, V. A. Vashchenko, S. O. Kolinko, and V. V. Tsybulin, "Research of the mechanism and development of the model of development of burning process of pyrotechnic mixes metal fuel + metal oxide at external thermal actions", *Vísnyk Cherkaskogo derzhavnogo tekhnologíchnogo uníversitetu*, no. 4, pp. 68-82, 2021 [in Ukrainian].

E. P. Kirichenko

Cherkasy Institute of Fire Safety,
Onoprienko st., 8, Cherkasy, 18034, Ukraine

TECHNIQUE FOR DETERMINATION OF CRITICAL VALUES OF PARAMETERS OF EXTERNAL THERMAL ACTIONS ON PYROTECHNICAL METAL OXIDE PRODUCTS IN OPERATING CONDITIONS

Ignition and subsequent fire and explosion destruction of products are preceded by external thermal actions to which they are exposed. This leads to premature ignition and explosive development of combustion of mixtures that are part of the products, and their subsequent flammable destruction. Therefore, techniques of prevention of forced premature fire-hazardous destruction of products in case of external thermal actions are of practical importance. They should be based on scientifically sound methods for determining critical values of thermal effects on products and technological parameters of mixtures, exceeding which leads to premature fire-hazardous destruction of products.

The aim of the work is to develop a technique for determining the critical values of external thermal influences and controlled technological parameters of mixtures, which lead to premature ignition and explosive combustion of pyrotechnic mixtures of magnesium, aluminum and metal oxides, and ultimately to flammable destruction of products.

A specialized software package of four standard packages of applications is formed. It consists of 80 individual software modules on modern numerical methods (nonlinear thermal conductivity

and diffusion, thermo-gas dynamics, chemical kinetics, mathematical processing of experimental data, etc.), which are focused on the personal computer that allows to solve in the mode of dialogue and real time typical engineering and technological problems of fire safety of various pyrotechnic mixtures.

A new scientifically sound technique based on the obtained mathematical, experimental and statistical models and specialized software in the form of four standard application packages has been developed. It allows in dialogue and real-time modes on the PC to form the necessary for practical use database of critical ranges of changes in the parameters of external thermal actions and controlled technological parameters of charges of mixtures, exceeding which leads to flammable destruction of products.

It is shown that practical use of the developed technique in the form of means of control and technological recommendations allows to increase fire safety of pyrotechnic products (lighting, signal, tracer cartridges and shells, incendiary means, etc.) in the conditions of external thermal actions in 2... 3 times.

Keywords: fire safety, pyrotechnic mixtures, ignition and combustion processes, powders of metal fuels and metal oxides.

- [0000-0002-9453-3144] **Ю. В. Сухацький**, канд. техн. наук,
e-mail: yurii.v.sukhatskyi@lpnu.ua
[0000-0001-9842-5864] **М. В. Шепіда**, канд. техн. наук,
e-mail: maryana_shepida@ukr.net
[0000-0003-1057-9928] **М. А. Созанський**, канд. хім. наук,
e-mail: martyn.a.sozanskyi@lpnu.ua
[0000-0002-3871-4063] **З. О. Знак**, д-р техн. наук, професор,
e-mail: znak_zo@ukr.net
[0000-0003-1931-1688] **С. В. Хом'як**, канд. хім. наук, с.н.с.
e-mail: semen.v.khomyak@lpnu.ua

Національний університет «Львівська політехніка»
вул. Степана Бандери, 12, м. Львів, 79013, Україна

КІНЕТИКА ОКИСНЮВАЛЬНОЇ ДЕГРАДАЦІЇ МЕТИЛОВОГО ОРАНЖЕВОГО З ВИКОРИСТАННЯМ ПРОЦЕСУ “СОНОПЕРОКСАТ”

Недоліками фотокаталітичних та Фентон-подібних передових процесів окиснення барвників є високі енерговитрати та довготривалість, що зумовлено утворенням тільки одного основного реакційного виду – гідроксильних радикалів. Для збільшення швидкості деградації азобарвників, зокрема метилового оранжевого (МО), вперше було запропоновано використовувати передовий процес окиснення “Сонопероксат” (ультразвук/ $\text{H}_2\text{O}_2/\text{KIO}_4$), під час реалізації якого формується комплекс реакційних видів, який включає як радикали (йодильні, перйодильні, гідроксильні), так і високоактивну сполуку молекулярного типу (синглетний кисень – $^1\text{O}_2$). Встановлено раціональні умови окиснювальної деградації МО: початкова концентрація МО у його водному розчині – $76,5 \cdot 10^{-6}$ моль/дм³; мольне співвідношення $\text{МО}:\text{H}_2\text{O}_2:\text{KIO}_4=1:50:10$; рН реакційного середовища – 3,00; температура середовища – 293 К; питома потужність ультразвукового кавітаційного оброблення – 68,0 Вт/дм³. Визначено константу швидкості окиснювальної деградації МО, яка за раціональних умов дорівнювала $2,883 \cdot 10^{-3} \text{ с}^{-1}$. Синергетичний ефект, зумовлений комбінацією дії ультразвукової кавітації та процесу $\text{H}_2\text{O}_2/\text{KIO}_4$, охарактеризовано значенням синергетичного коефіцієнта, який дорівнював 1,62. На основі результатів UV/Vis- та ІЧ-спектроскопії підтверджено руйнування азозв'язку у молекулах МО та деградацію барвника. За результатами ідентифікації функціональних груп інтермедіатів методом ІЧ-спектроскопії запропоновано можливий шлях окиснювальної деградації МО з використанням процесу “Сонопероксат”, який послідовно включає розривання азозв'язку, деметилювання первинних амінів, деамінування амінопохідних бензену, розкриття бензенового кільця з утворенням лінійних сполук та мінералізацію останніх.

Ключові слова: азобарвники, окиснювальна деградація, константа швидкості, ультразвукова кавітація, періодат, UV/Vis-спектроскопія, ІЧ-спектроскопія.

Вступ. Потрапляння високозабарвлених стічних вод виробництв барвників, текстильної, целюлозно-паперової, фармацевтичної, харчової, переробної та інших галузей промисловості у природні водойми суттєво зменшує проникну здатність світлового випромінювання. Це зумовлює значне зниження потенціалу рослин, які входять до складу водних екосистем, до фотосинтезу і, відповідно, зменшення вмісту розчиненого кисню у воді [1]. Акумуляція синтетичних барвників, канцерогенна і мутагенна дія яких є загальновідомою, у водних середовищах призводить до їх неминучого залучення у трофічні ланцюги.

Результатом стає загибель кількох рівнів організмів таких ланцюгів.

Традиційні технологічні процеси (адсорбція, йонний обмін, мембранне фільтрування тощо) є високовартісними, характеризуються накопиченням шкідливих осадів, які лише частково піддаються рекуперації та утилізації. Тому на сьогодні увага дослідників спрямована на розроблення нових технологій очищення вод від барвників, які б передбачали не тільки вилучення барвника з водного середовища, а й також його мінералізацію до вуглекислого газу, води та інших нешкідли-

вих продуктів. До таких сучасних технологій належать передові процеси окиснення.

Азобарвники становлять понад 50 % від усієї номенклатури синтетичних барвників [2]. Найчастіше для деградації азобарвників використовують такі передові процеси окиснення, як фотокаталіз [3-12], процес Фентона та Фентон-подібні процеси [13, 14].

Внаслідок додавання 50 мг фотокаталізатора ZnO/CuO до 50 см³ водного розчину конго червоного (концентрація барвника – 50 мг/дм³) впродовж 3000 с дії видимого випромінювання було досягнуто ступеня деградації барвника 91 % [10]. Початкове значення рН водного розчину конго червоного дорівнювало 5,6.

Застосування процесу електро-Фентон впродовж 5400 с зумовило повне знебарвлення водного розчину метилового оранжевого [13]. Умови окиснювальної деградації метилового оранжевого були такими: концентрація барвника – 75 мг/дм³; концентрація H₂O₂ – 5·10⁻² моль/дм³; концентрація йонів Fe²⁺ – 10⁻⁴ моль/дм³; рН – 2,9; густина струму – 15 мА/см². При цьому зафіксували зменшення значення хімічного споживання кисню розчину на 95 %.

Основними недоліками зазначених вище передових процесів окиснення є високі енерговитрати та довготривалість. Довготривалість деградації барвників з використанням фотокаталітичних та Фентон-подібних процесів окиснення визначається, насамперед, утворенням тільки одного основного реакційного виду – гідроксильних радикалів. Тому актуальним є розроблення нових передових процесів окиснення, під час реалізації яких генерується комплекс реакційних видів – як радикалів, так і нерадикальних видів. Це дозволить суттєво збільшити швидкість деградації азобарвників.

Науково-дослідною групою під керівництвом професора О. Hamdaoui було розроблено новий передовий процес окиснення – процес “Пероксат” (H₂O₂/KIO₄), у якому перйодати активують гідрогену пероксидом [15, 16]. Як наслідок, формується комплекс окисників, який поєднує як радикали (йодильні, перйодильні, гідроксильні), так і високоактивну сполуку молекулярного типу (синглетний кисень –¹O₂). Дослідники використали процес “Пероксат” для інтенсивної деградації тіазинового барвника Toluidine blue (концентрація барвника – 10 мг/дм³). Так, вже впродовж 60 с за рН=5,4, температури реакційного середо-

вища 298 К і мольного співвідношення H₂O₂:IO₄⁻=10:1 було досягнуто ступеня деградації барвника 98 %, а концентрація загального органічного вуглецю при цьому зменшилась на 73 % [16].

Поєднання кількох стратегій активації перйодатів, наприклад активації гідрогену пероксидом та ультразвуковими (УЗ) коливаннями, зумовило істотне збільшення константи швидкості знебарвлення тіазинового барвника метиленового синього, порівняно із використанням для активації перйодатів тільки одного фактора. Так, константа швидкості знебарвлення метиленового синього у випадку: активації перйодатів УЗ-коливаннями (УЗ/KIO₄) дорівнювала 3,781·10⁻⁵ с⁻¹, активації гідрогену пероксидом (H₂O₂/KIO₄) – 2,201·10⁻³ с⁻¹, комбінованої стратегії активації (УЗ/H₂O₂/KIO₄) – 7,369·10⁻³ с⁻¹ [17]. Синергетичний коефіцієнт процесу “Сонопероксат” (УЗ/H₂O₂/KIO₄), розрахований на основі констант швидкості знебарвлення, дорівнював 3,25.

Перспективним напрямом подальших досліджень є апробація процесу “Сонопероксат” для деградації не тільки представників класу тіазинових барвників, але й барвників інших класів – наприклад, азобарвників.

Мета та завдання дослідження. Метою роботи було дослідження впливу низки факторів (рН реакційного середовища, мольного співвідношення барвник:H₂O₂/KIO₄, питомої потужності ультразвукового кавітаційного оброблення, початкової концентрації барвника у його водному розчині, температури середовища) на кінетику окиснювальної деградації моноазобарвника метилового оранжевого (МО) з використанням процесу “Сонопероксат”.

Для досягнення зазначеної мети необхідно вирішити такі завдання:

1) визначити кінетичні параметри (константу швидкості, ефективну енергію активації) окиснювальної деградації МО з використанням процесу “Сонопероксат”;

2) розрахувати синергетичний коефіцієнт процесу УЗ/H₂O₂/KIO₄ для деградації МО;

3) за результатами ІЧ-спектроскопії реакційного середовища ідентифікувати функціональні групи інтермедіатів і запропонувати можливий шлях окиснювальної деградації МО з використанням процесу “Сонопероксат”.

Матеріали та методи досліджень. Основними елементами установки для дослі-

джені кінетики окиснювальної деградації МО були: скляний реактор ($V=250\text{ см}^3$); УЗ-випромінювач магнітострикційного типу із блоком живлення “Ultrasonic Disintegrator UD-20” (Польща) зі змінною питомою потужністю ультразвукового кавітаційного оброблення ($53,3$; $68,0$ і $83,3\text{ Вт/дм}^3$) та діаметром робочої частини концентратора УЗ (частота УЗ-коливань – 22 кГц) 15 мм і глибиною його занурення в розчин $10\pm 1\text{ мм}$; термостат CORIO C-BT5 (JULABO GmbH) для створення ізотермічних умов у реакційному середовищі; магнітна мішалка. Реактор поміщали у виносну термостатну камеру з термометром, якою було оснащено термостат.

150 см^3 водного розчину МО з відповідною концентрацією барвника 15 мг/дм^3 ($45,9\cdot 10^{-6}\text{ моль/дм}^3$), 25 мг/дм^3 ($76,5\cdot 10^{-6}\text{ моль/дм}^3$) або 35 мг/дм^3 ($107\cdot 10^{-6}\text{ моль/дм}^3$) готували з використанням дистильованої води за постійного перемішування. Приготованим розчином заповнювали скляний реактор (коефіцієнт заповнення реактора – 60%). Реакції окиснення ініціювали послідовним дозуванням у водний розчин МО певного об'єму 35% -го водного розчину гідрогену пероксиду (Merck, Німеччина) та певної кількості порошку калію перйодату (Merck, Німеччина) з їх подальшим перемішуванням. рН реакційного середовища (у діапазоні від $3,00$ до $11,00$) регулювали хлоридною кислотою або натрію гідроксидом. Для вимірювання рН використовували рН-метр ADWA AD1200 ATC (Угорщина). Ізотермічні умови забезпечували шляхом термостатування реакційного середовища у скляному реакторі.

Комплексом експериментальних досліджень було передбачено аналіз впливу таких факторів на кінетику окиснювальної деградації МО: 1) рН реакційного середовища; 2) мольного співвідношення $\text{МО}:\text{H}_2\text{O}_2:\text{KIO}_4$, яке змінювали у межах від $1:25:5$ до $1:100:20$; 3) питомої потужності ультразвукового кавітаційного оброблення; 4) початкової концентрації барвника у його водному розчині; 5) температури середовища.

Через задані проміжки часу відбирали проби реакційного середовища, які аналізували з використанням UV-Vis-спектрофотометра UV-3100PC (Shanghai Mapada Instruments Co., Ltd.) з набором кварцових кювет товщиною 1 см у діапазоні довжин хвиль від 200 до 1100 нм . За рН=3,00 максимум поглинання МО відповідає довжині хвилі 500 нм , а за рН=7,00; 11,00 – довжині хвилі 464 нм . Бато-

хромне (червоне) зміщення максимуму поглинання МО у кислому середовищі від 464 до 500 нм обумовлене протонуванням барвника за азогрупою з утворенням йона азонію [18]. Саме ця реакційна форма є доміантною за низьких значень рН. Концентрацію МО обчислювали на основі попередньо побудованих калібрувальних кривих.

Синергетичний коефіцієнт процесу (f) $\text{УЗ}/\text{H}_2\text{O}_2/\text{KIO}_4$ для деградації МО розраховували на основі значень констант швидкості

$$f = \frac{k_{\text{УЗ}/\text{H}_2\text{O}_2/\text{KIO}_4}}{k_{\text{УЗ}/\text{H}_2\text{O}_2} + k_{\text{H}_2\text{O}_2/\text{KIO}_4} + k_{\text{УЗ}/\text{KIO}_4}}. \quad (1)$$

Функціональні групи інтермедіатів окиснювальної деградації МО ідентифікували на основі результатів ІЧ-спектроскопії. ІЧ-спектри розчину МО (спектральний діапазон – $4000\ldots 400\text{ см}^{-1}$) під час його окиснювальної деградації отримували з використанням ІЧ-Фур'є спектрометра Spectrum Two (Perkin Elmer), оснащеного універсальною приставкою порушеного повного внутрішнього відбиття.

Результати досліджень та їх обговорення. Найінтенсивніше деградація МО відбувалась у кислому середовищі, за рН=3,00 (рисунк 1).

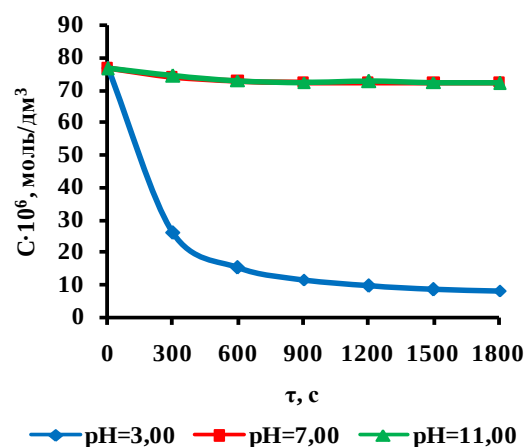
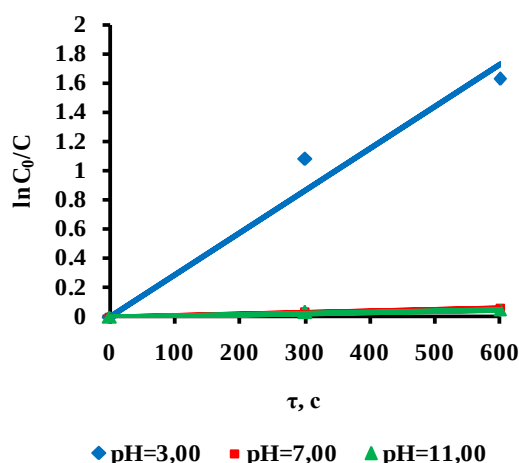


Рисунок 1 – Залежність концентрації МО в реакційному середовищі від тривалості окиснювальної деградації з використанням процесу “Сонопероксат” за різних значень рН середовища (3,00...11,00):
(експериментальні умови:
 $C_0(\text{МО})=76,5\cdot 10^{-6}\text{ моль/дм}^3$, мольне співвідношення $\text{МО}:\text{H}_2\text{O}_2:\text{KIO}_4=1:50:10$, $T=293\text{ К}$, питома потужність ультразвукового кавітаційного оброблення – $68,0\text{ Вт/дм}^3$)

Впродовж 900 с комбінованого ультразвукового та реагентного оброблення концентрація МО в реакційному середовищі зменшилась від $76,5 \cdot 10^{-6}$ до $11,5 \cdot 10^{-6}$ моль/дм³ (рисунки 1), що еквівалентно ступеню окиснювальної деградації 84,9 %. Збільшення тривалості оброблення до 1800 с призводило до зростання ступеня деградації до 89,4 %.

Результати, які наведено на рисунку 2, вказують на те, що деградація МО з використанням процесу “Сонопероксат” протікає як реакція псевдопершого порядку. Зі зростанням рН середовища від 3,00 до 7,00 внаслідок застосування процесу $U_3/H_2O_2/KIO_4$ константа швидкості окиснювальної деградації МО зменшувалась у 31,8 рази – від $2,883 \cdot 10^{-3}$ до $9,067 \cdot 10^{-5}$ с⁻¹ (рисунки 2). Подальше зростання рН реакційного середовища до 11,00 призводило до зменшення константи швидкості деградації МО на 7,2 % – від $9,067 \cdot 10^{-5}$ до $8,455 \cdot 10^{-5}$ с⁻¹.

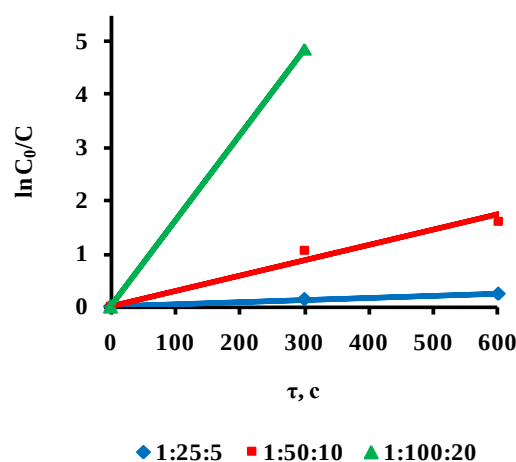


Рисунки 2 – Лінійні анаморфози кінетичних кривих окиснювальної деградації МО (вплив рН реакційного середовища) (експериментальні умови: $C_0(МО) = 76,5 \cdot 10^{-6}$ моль/дм³, мольне співвідношення $МО:H_2O_2:KIO_4 = 1:50:10$, $T = 293$ К, питома потужність ультразвукового кавітаційного оброблення – $68,0$ Вт/дм³)

Суттєве зменшення ефективності (від 89,4 до 5,6 %) та інтенсивності деградації МО в лужному середовищі (рН=11,00) зумовлене утворенням карбонатних радикалів – слабших окисників, порівняно з гідроксильними радикалами та синглетним киснем, що є основними реакційними видами передового процесу окиснення “Сонопероксат” [17]. Утворення карбонатних радикалів – результат поглинання гідроксильних радикалів та супероксидно-

го аніона-радикала, який є попередником синглетного кисню, гідрогенкарбонат- і карбонат-йонами. Ці йони формуються з вуглекислого газу, що є продуктом мінералізації МО.

Збільшення мольного співвідношення $МО:H_2O_2:KIO_4$ від 1:25:5 до 1:50:10 призводило до збільшення константи швидкості окиснювальної деградації МО у 7,3 рази – від $3,971 \cdot 10^{-4}$ до $2,883 \cdot 10^{-3}$ с⁻¹ (рисунки 3). Подальше збільшення мольного співвідношення $МО:H_2O_2:KIO_4$ до 1:100:20 зумовлювало зростання константи швидкості в 5,6 рази – від $2,883 \cdot 10^{-3}$ до $1,622 \cdot 10^{-2}$ с⁻¹.

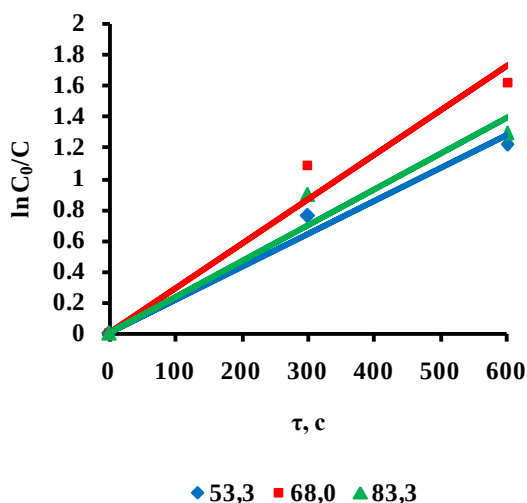


Рисунки 3 – Лінійні анаморфози кінетичних кривих окиснювальної деградації МО (вплив мольного співвідношення $МО:H_2O_2:KIO_4$) (експериментальні умови: $C_0(МО) = 76,5 \cdot 10^{-6}$ моль/дм³, рН=3,00, $T = 293$ К, питома потужність ультразвукового кавітаційного оброблення – $68,0$ Вт/дм³)

При цьому впродовж 300 с оброблення досягнули ступеня окиснювальної деградації МО: 12,4 % (за мольного співвідношення $МО:H_2O_2:KIO_4 = 1:25:5$); 66,0 % (за мольного співвідношення $МО:H_2O_2:KIO_4 = 1:50:10$) і 99,2 % (за мольного співвідношення $МО:H_2O_2:KIO_4 = 1:100:20$). Але для зменшення собівартості очищення за рахунок зменшення витрат окисників (гідрогену пероксиду та калію перйодату) було запропоновано збільшити тривалість процесу від 300 до 1800 с. У такому випадку ступінь деградації МО за мольного співвідношення $МО:H_2O_2:KIO_4 = 1:50:10$ зростав від 66,0 до 89,4 %.

Залежність константи швидкості окиснювальної деградації МО від питомої потужності ультразвукового кавітаційного оброб-

лення мала екстремальний характер із максимумом за 68,0 Вт/дм³. Так, зі збільшенням питомої потужності від 53,3 до 68,0 Вт/дм³ константа швидкості зростала на 35,4 % – від $2,130 \cdot 10^{-3}$ до $2,883 \cdot 10^{-3}$ с⁻¹ (рисунк 4). Подальше збільшення питомої потужності до 83,3 Вт/дм³ призводило до зменшення константи швидкості деградації МО на 19,4 % – від $2,883 \cdot 10^{-3}$ до $2,323 \cdot 10^{-3}$ с⁻¹.

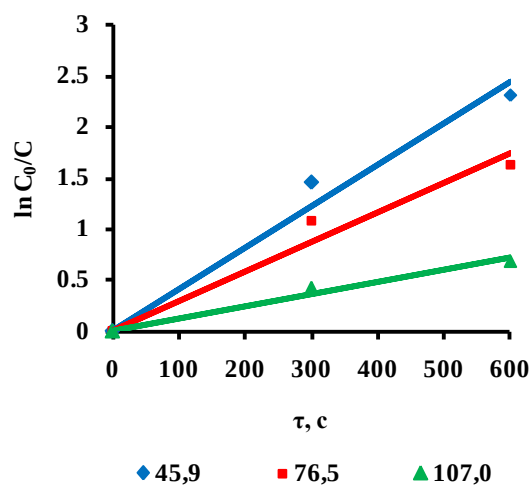


Рисунк 4 – Лінійні анаморфози кінетичних кривих окиснювальної деградації МО (вплив питомої потужності ультразвукового кавітаційного оброблення, Вт/дм³):
(експериментальні умови:
 $C_0(\text{МО}) = 76,5 \cdot 10^{-6}$ моль/дм³,
мольне співвідношення $\text{МО}:\text{H}_2\text{O}_2:\text{KIO}_4 = 1:50:10$,
 $\text{pH} = 3,00$, $T = 293$ К)

Наведене вище можна пояснити виникненням у рідкій фазі за значень питомої потужності понад 68,0 Вт/дм³ явища “виродженої кавітації” – генерування дуже дрібних кавітаційних бульбашок та їх агломератів, стійких до дії механічних, хімічних і термічних факторів [19]. Надзвичайна стабільність таких бульбашок зменшує ймовірність їх сплескування і, відповідно, ефективність та інтенсивність окиснювальної деградації МО.

Зі збільшенням початкової концентрації забруднювача (моноазобарвника МО) у реакційному середовищі за постійних витрат окисників (гідрогену пероксиду – $3,825 \cdot 10^{-3}$ моль/дм³; калію періодату – $0,765 \cdot 10^{-3}$ моль/дм³) константа швидкості деградації МО зменшувалась. Так, за початкової концентрації МО в реакційному середовищі $45,9 \cdot 10^{-6}$ моль/дм³ константа швидкості дегра-

дації МО дорівнювала $4,048 \cdot 10^{-3}$ с⁻¹, за концентрації $76,5 \cdot 10^{-6}$ моль/дм³ – $2,883 \cdot 10^{-3}$ с⁻¹, а за концентрації $107,0 \cdot 10^{-6}$ моль/дм³ – $1,192 \cdot 10^{-3}$ с⁻¹ (рисунк 5).



Рисунк 5 – Лінійні анаморфози кінетичних кривих окиснювальної деградації МО (вплив початкової концентрації МО в реакційному середовищі, $C_0(\text{МО}) \cdot 10^6$ моль/дм³):
(експериментальні умови:
 $C_0(\text{H}_2\text{O}_2) = 3,825$ ммоль/дм³,
 $C_0(\text{KIO}_4) = 0,765$ ммоль/дм³, $\text{pH} = 3,00$, $T = 293$ К,
питома потужність ультразвукового кавітаційного оброблення – 68,0 Вт/дм³)

Зі збільшенням температури реакційного середовища константа швидкості окиснювальної деградації МО монотонно зменшувалась. Так, внаслідок підвищення температури від 293 до 308 К константа швидкості зменшилась в 1,5 рази – від $2,883 \cdot 10^{-3}$ до $1,918 \cdot 10^{-3}$ с⁻¹ (рисунк 6). Подальше підвищення температури до 323 К зумовлювало зменшення константи швидкості в 1,4 рази – від $1,918 \cdot 10^{-3}$ до $1,373 \cdot 10^{-3}$ с⁻¹. Такий характер залежності константи швидкості деградації МО від температури реакційного середовища можна пояснити збільшенням ступеня дисоціації гідрогену пероксиду за кислотним механізмом з утворенням H^+ і HO_2^- . Це призводить до непродуктивної витрати гідрогену пероксиду і, відповідно, зменшення ступеня його залученості до генерування сильніших окисників, таких як гідроксильні, йодильні радикали та синглетний кисень за механізмом процесу $\text{U}_3/\text{H}_2\text{O}_2/\text{KIO}_4$ [17].

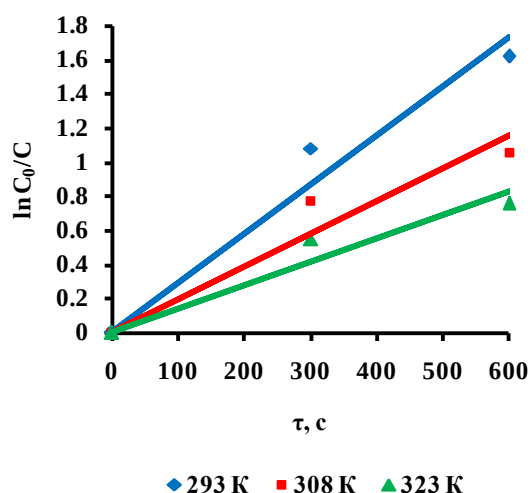


Рисунок 6 – Лінійні анаморфози кінетичних кривих окиснювальної деградації МО (вплив температури реакційного середовища): (експериментальні умови: $C_0(\text{МО})=76,5 \cdot 10^{-6}$ моль/дм³, мольне співвідношення $\text{МО}:\text{H}_2\text{O}_2:\text{KIO}_4=1:50:10$, $\text{pH}=3,00$, питома потужність ультразвукового кавітаційного оброблення – $68,0$ Вт/дм³)

На підставі залежності, яку наведено на рисунку 7, розраховано значення ефективної енергії активації (E_a) окиснювальної деградації МО із використанням процесу “Сонопероксат”

$$E_a = -Rt\alpha = -8,314 \cdot 2342,112 \approx -19472 \text{ Дж/моль}, \quad (2)$$

де R – універсальна газова стала, Дж/(моль·К) ($R=8,314$ Дж/(моль·К)); $\text{tg } \alpha$ – тангенс кута нахилу прямої $\ln k=f(1/T)$ до осі абсцис, К.

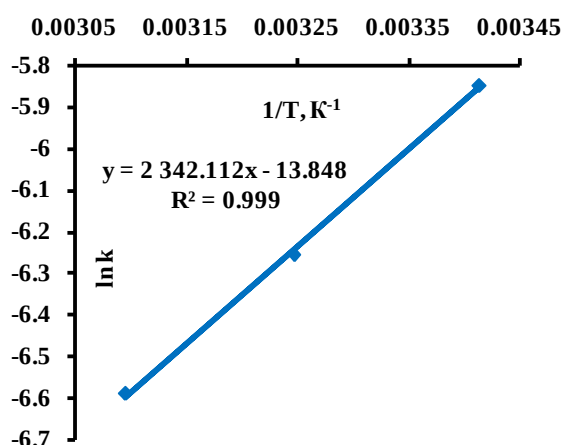


Рисунок 7 – Залежність $\ln k$ від $1/T$

Від’ємне значення ефективної енергії активації окиснювальної деградації МО із використанням передового процесу окиснення “Сонопероксат” пов’язане з багатостадійністю цього процесу. Так, запропонований механізм процесу “Сонопероксат” включає 32 реакції за участю радикалів, зокрема гідрогенпероксидних ($\text{HO}_2^\bullet$) та гідроксильних ($^\bullet\text{OH}$) [17]. Відповідно до даних, які представили Mozurkewich і Benson у публікації [20], до реакцій із від’ємними значеннями енергії активації належать взаємодія гідроксильних радикалів з нітратною або пероксонітратною кислотою, а також диспропорціонування гідрогенпероксидних радикалів. Від’ємні значення енергії активації цих реакцій автори пояснили утворенням стабільних проміжних комплексів. Ймовірно, окремі стадії процесу “Сонопероксат” за участю гідроксильних та гідрогенпероксидних радикалів є лімітуючими і саме вони визначають від’ємне значення ефективної енергії активації окиснювальної деградації МО.

Встановлено, що ступінь деградації МО з використанням процесів $\text{УЗ}/\text{H}_2\text{O}_2/\text{KIO}_4$, $\text{H}_2\text{O}_2/\text{KIO}_4$, $\text{УЗ}/\text{KIO}_4$ та $\text{УЗ}/\text{H}_2\text{O}_2$ впродовж 1800 с (рисунк 8) дорівнював відповідно 89,4; 83,7; 5,3 та 4,4 %.

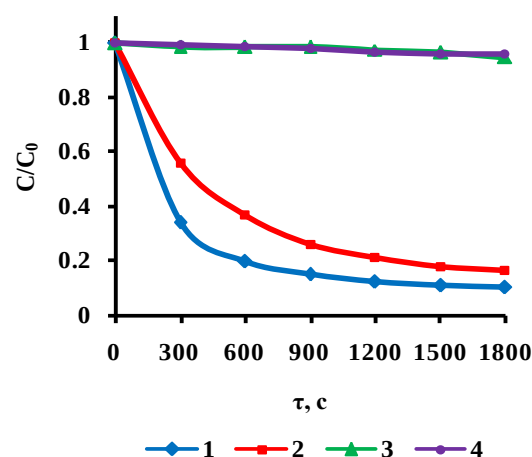


Рисунок 8 – Кінетичні криві деградації МО з використанням процесів: 1 – $\text{УЗ}/\text{H}_2\text{O}_2/\text{KIO}_4$; 2 – $\text{H}_2\text{O}_2/\text{KIO}_4$; 3 – $\text{УЗ}/\text{KIO}_4$; 4 – $\text{УЗ}/\text{H}_2\text{O}_2$ (експериментальні умови: $C_0(\text{МО})=76,5 \cdot 10^{-6}$ моль/дм³, мольне співвідношення $\text{МО}:\text{H}_2\text{O}_2:\text{KIO}_4=1:50:10$, $\text{pH}=3,00$, $T=293$ К, питома потужність ультразвукового кавітаційного оброблення – $68,0$ Вт/дм³)

Константа швидкості окиснювальної деградації МО з використанням процесів $\text{УЗ}/\text{H}_2\text{O}_2/\text{KIO}_4$, $\text{H}_2\text{O}_2/\text{KIO}_4$, $\text{УЗ}/\text{KIO}_4$ та $\text{УЗ}/\text{H}_2\text{O}_2$ (рисунок 9) дорівнювала відповідно $2,883 \cdot 10^{-3}$; $1,721 \cdot 10^{-3}$; $3,250 \cdot 10^{-5}$ і $2,712 \cdot 10^{-5} \text{ с}^{-1}$.

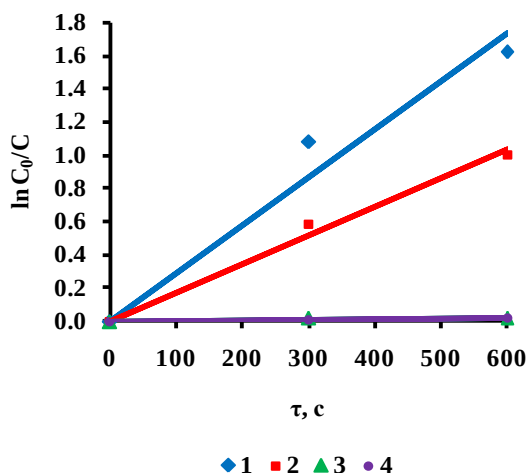


Рисунок 9 – Лінійні анаморфози кінетичних кривих деградації МО з використанням процесів: 1 – $\text{УЗ}/\text{H}_2\text{O}_2/\text{KIO}_4$; 2 – $\text{H}_2\text{O}_2/\text{KIO}_4$; 3 – $\text{УЗ}/\text{KIO}_4$; 4 – $\text{УЗ}/\text{H}_2\text{O}_2$

(експериментальні умови:

$C_0(\text{МО})=76,5 \cdot 10^{-6} \text{ моль/дм}^3$, мольне співвідношення $\text{МО}:\text{H}_2\text{O}_2:\text{KIO}_4=1:50:10$, $\text{pH}=3,00$, $T=293 \text{ К}$, питома потужність ультразвукового кавітаційного оброблення – $68,0 \text{ Вт/дм}^3$)

Синергетичний ефект, зумовлений комбінацією дії ультразвукової кавітації та процесу $\text{H}_2\text{O}_2/\text{KIO}_4$, охарактеризовано значенням синергетичного коефіцієнта (f), який обчислювали на основі значень констант швидкості деградації МО для окремих процесів

$$f = \frac{2,883 \cdot 10^{-3}}{2,712 \cdot 10^{-5} + 1,721 \cdot 10^{-3} + 3,250 \cdot 10^{-5}} = 1,62. \quad (3)$$

Отже, запропонований комбінований передовий процес окиснення ($\text{УЗ}/\text{H}_2\text{O}_2/\text{KIO}_4$) є значно інтенсивнішим та ефективнішим, порівняно з окремими процесами, які використали для його розроблення. Це підтверджено значеннями коефіцієнта синергізму та ступеня окиснювальної деградації МО.

На рисунку 10 наведено UV/Vis-спектри окиснювальної деградації МО з використанням процесу $\text{УЗ}/\text{H}_2\text{O}_2/\text{KIO}_4$. Для МО характерні максимуми (піки) поглинання за довжини хвилі 500 нм (у кислому середовищі) або 464 нм (у нейтральному та лужному середовищі). Зміщення максимуму поглинання МО у кислому

середовищі в область більших довжин хвиль (від 464 до 500 нм), так зване батохромне (червоне) зміщення, зумовлене протонуванням барвника за азогрупою з утворенням йона азонію [18]. Піки за довжин хвиль 464 і 500 нм відповідають $n \rightarrow \pi^*$ переходам у хромофорних $\text{N}=\text{N}$ і $\text{C}=\text{N}$ групах [21].

Під час окиснювальної деградації МО з використанням процесу $\text{УЗ}/\text{H}_2\text{O}_2/\text{KIO}_4$ спостерігали гіпсохромне (синє) зміщення максимуму поглинання від 500 до 464 нм (рисунок 10), що було зумовлено депротонуванням МО за азогрупою. Крім того, було зафіксовано суттєве зменшення інтенсивності поглинання (гіпохромний ефект) як за довжини хвилі 500 нм, так і за 464 нм. Такі зміни UV/Vis-спектрів свідчать про руйнування хромофорних груп барвника – відповідно, $\text{C}=\text{N}$ та азогрупи ($\text{N}=\text{N}$) [18, 21]. Зазначене було підтверджено знебарвленням реакційного середовища. Зменшення інтенсивності поглинання у діапазоні довжин хвиль 250...300 нм, що відповідають $\pi \rightarrow \pi^*$ переходам у бензенових (ароматичних) кільцях [21, 22], внаслідок застосування процесу “Сонопероксат” щонайменше впродовж 900 с вказує на розкриття ароматичних (бензенових) кілець з утворенням лінійних продуктів. Це опосередковано свідчить про мінералізацію барвника.

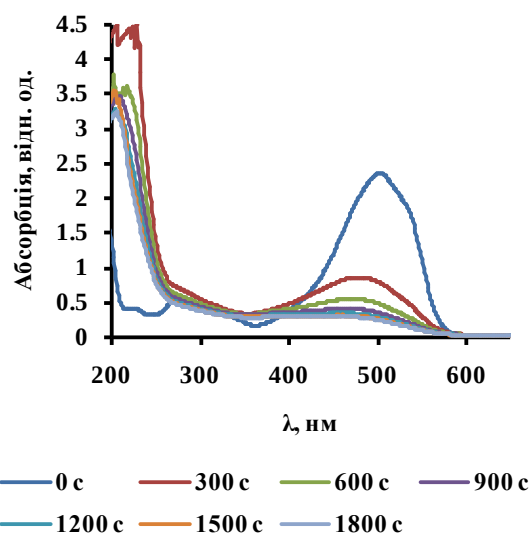


Рисунок 10 – UV/Vis-спектри окиснювальної деградації МО з використанням процесу $\text{УЗ}/\text{H}_2\text{O}_2/\text{KIO}_4$

(експериментальні умови:

$C_0(\text{МО})=76,5 \cdot 10^{-6} \text{ моль/дм}^3$, мольне співвідношення $\text{МО}:\text{H}_2\text{O}_2:\text{KIO}_4=1:50:10$, $\text{pH}=3,00$, $T=293 \text{ К}$, питома потужність ультразвукового кавітаційного оброблення – $68,0 \text{ Вт/дм}^3$)

На рисунку 11 наведено результати ІЧ-спектроскопії для окиснювальної деградації МО з використанням процесу $УЗ/H_2O_2/KIO_4$.

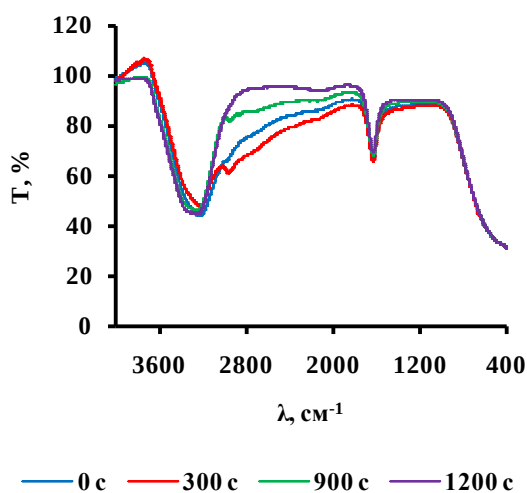


Рисунок 11 – ІЧ-спектри окиснювальної деградації МО з використанням процесу $УЗ/H_2O_2/KIO_4$

(експериментальні умови:
 $C_0(MO) = 76,5 \cdot 10^{-6}$ моль/дм³, мольне співвідношення $MO:H_2O_2:KIO_4 = 1:50:10$, $pH = 3,00$, $T = 293$ К, питома потужність ультразвукового кавітаційного оброблення – $68,0$ Вт/дм³)

Аналіз ІЧ-спектру початкового водного розчину МО до його сумісного кавітаційного та реагентного оброблення (0 с) свідчить про наявність піків за 3222 см^{-1} , що вказує на валентні коливання N–H у вторинних амідах, і 1635 см^{-1} , що вказує на валентні коливання C=O в амідах і C=C в алкенах. Після 300 с використання процесу $УЗ/H_2O_2/KIO_4$, крім піків за 3222 см^{-1} і 1635 см^{-1} (інтенсивність піку дещо зростала, порівняно з ІЧ-спектром необробленого водного розчину МО), виникав пік за 2967 см^{-1} , що відповідав валентним асиметричним коливанням C–H в алканах. Збільшення тривалості очищення до 900 с призводило до зникнення піку за 3222 см^{-1} . Це підтверджує розривання азозв'язку [1, 23]. Крім того, на ІЧ-спектрі реакційного середовища, обробленого впродовж 900 с, з'являвся пік за 3257 см^{-1} , що відповідав валентним коливанням N–H первинних ароматичних амінів [1]. Наявність піку за 2956 см^{-1} вказувало на валентні асиметричні коливання C–H в алканах, а піку за 1635 см^{-1} – на валентні коливання C=O в амідах і C=C в алкенах. Після застосування процесу “Сонопероксат” впродовж 1200 с на ІЧ-спектрі зникали піки за 3257 і

2956 см^{-1} , а інтенсивність піку за 1635 см^{-1} , що відповідав валентним коливанням C=O в амідах і C=C в алкенах, дещо зменшувалась, що вказувало на їх мінералізацію. Натомість з'являвся пік за 3265 см^{-1} , що відповідав валентним коливанням =C–H в алкенах. Зазначене вище свідчить, що під час окиснювальної деградації МО з використанням процесу $УЗ/H_2O_2/KIO_4$ послідовно відбуваються розривання азозв'язку, деметилювання первинних амінів, деамінування амінопохідних бензену, розкриття бензенових кілець з утворенням лінійних сполук та мінералізація останніх. Необхідно зазначити взаємоузгодженість результатів UV/Vis- та ІЧ-спектроскопії.

Висновки. Наукова новизна роботи полягає у тому, що вперше було визначено кінетичні параметри (константу швидкості, ефективну енергію активації) окиснювальної деградації моноазобарвника МО з використанням процесу $УЗ/H_2O_2/KIO_4$.

Встановлено раціональні умови окиснювальної деградації МО: початкова концентрація МО у його водному розчині – $76,5 \cdot 10^{-6}$ моль/дм³; мольне співвідношення $MO:H_2O_2:KIO_4 = 1:50:10$; pH реакційного середовища – 3,00; температура середовища – 293 К; питома потужність ультразвукового кавітаційного оброблення – $68,0$ Вт/дм³. За таких умов константа швидкості дорівнювала $2,883 \cdot 10^{-3}\text{ с}^{-1}$. Синергетичний ефект, зумовлений комбінацією дії ультразвукової кавітації та процесу H_2O_2/KIO_4 , було охарактеризовано значенням синергетичного коефіцієнта, який дорівнював 1,62. На основі результатів UV/Vis- та ІЧ-спектроскопії підтверджено руйнування азозв'язку у молекулах МО та деградацію барвника. Ідентифікація функціональних груп інтермедіатів методом ІЧ-спектроскопії дала змогу запропонувати можливий шлях окиснювальної деградації МО з використанням процесу “Сонопероксат”, який послідовно включає розривання азозв'язку, деметилювання первинних амінів, деамінування амінопохідних бензену, розкриття бензенових кілець з утворенням лінійних сполук та мінералізацію останніх.

Результати кінетичних досліджень окиснювальної деградації азобарвників з використанням процесу $УЗ/H_2O_2/KIO_4$ стануть підґрунтям для реалізації цього передового процесу окиснення у промислових масштабах.

Перспективним напрямом подальших досліджень є комбінація ультразвукової акти-

вації періодатів з активацією сполуками d-елементів.

Роботу виконано за підтримки Міністерства освіти і науки України у межах проєкту наукової роботи молодих вчених "Передові процеси окиснення, зокрема нанокаталітичного, в основі кавітаційних технологій очищення водних середовищ від резистентних N-заміщених органічних сполук" (номер державної реєстрації 0122U000790).

Автори висловлюють подяку колективу ЦККНО "Лабораторія перспективних технологій створення та фізико-хімічного аналізу нових речовин і функціональних матеріалів" при Національному університеті "Львівська політехніка" за ефективну співпрацю та всебічну підтримку під час виконання експериментальних досліджень.

Список використаних джерел / References

- [1] H. N. Altayb, B. Kouidhi, O. A. S. Baothman, J. A. Abdulhakim, L. Ayed, M. Hager, and K. Chaieb, "Mathematical modeling and optimization by the application of full factorial design and response surface methodology approach for decolourization of dyes by a newly isolated *Photobacterium ganghwense*", *Journal of Water Process Engineering*, vol. 44, p. 102429, 2021.
doi: 10.1016/j.jwpe.2021.102429.
- [2] Y. Sukhatskiy, Z. Znak, O. Zin, and D. Chupinskiy, "Ultrasonic cavitation in wastewater treatment from azo dye methyl orange", *Chemistry & Chemical Technology*, vol. 15, no. 2, pp. 284-290, 2021.
doi: 10.23939/chcht15.02.284.
- [3] M. Saeed, S. Adeel, H. Abdur-raoof, M. U. A. Mansha, A. Ahmad, and M. Amjed, "ZnO catalyzed degradation of methyl orange in aqueous medium", *Chiang Mai Journal of Science*, vol. 44 (4), pp. 1646-1653, 2017.
doi: 10.13140/rg.2.2.17237.01762.
- [4] R. A. Putri, S. Safni, N. Jamarun, and U. Septiani, "Kinetics study and degradation pathway of methyl orange photodegradation in the presence of C-N-codoped TiO₂ catalyst", *Egyptian Journal of Chemistry*, vol. 62, pp. 563-575, 2019.
doi: 10.21608/ejchem.2019.14543.1883.
- [5] Z. Junbo, M. Di, Z. Hong, L. An, L. M. Jiao, H. Shengtian, and L. Jianzhang, "Photocatalytic decolorization of methyl orange solution with potassium peroxydisulfate", *Central European Journal of Chemistry*, vol. 6 (2), pp. 245-252, 2008.
doi: 10.2478/s11532-008-0016-5.
- [6] P. S. Harikumar, L. Joseph, and A. Dhanya, "Photocatalytic degradation of textile dyes by hydrogel supported titanium dioxide nanoparticles", *Journal of Environmental Engineering & Ecological Science*, vol. 2, p. 2, 2013.
doi: 10.7243/2050-1323-2-2.
- [7] S. A. Bhat, F. Zafar, A. H. Mondal, A. Kareem, A. Z. Mirza, S. Khan, A. Mohammad, Q. M. R. Haq, and N. Nishat, "Photocatalytic degradation of carcinogenic Congo red dye in aqueous solution, antioxidant activity and bactericidal effect of NiO nanoparticles", *Journal of the Iranian Chemical Society*, vol. 17 (1), pp. 215-227, 2020.
doi: 10.1007/s13738-019-01767-3.
- [8] V. K. Landge, C.-M. Huang, V. S. Hakke, S. H. Sonawane, S. Manickam, and M.-C. Hsieh, "Solar-energy-driven Cu-ZnO/TiO₂ nanocomposite photocatalyst for the rapid degradation of Congo red azo dye", *Catalysts*, vol. 12, p. 605, 2022.
doi: 10.3390/catal12060605.
- [9] M. Said, W. T. Rizki, W. R. Asri, D. Desnelli, A. Rachmat, and P. L. Hariyani, "SnO₂-Fe₃O₄ nanocomposites for the photodegradation of the Congo red dye", *Heliyon*, vol. 8 (4), p. e09204, 2022.
doi: 10.1016/j.heliyon.2022.e09204.
- [10] G. Hitkari, P. Ghoshdary, V. Kumar, S. Singh, and A. Motghare, "Potential of copper-zinc oxide nanocomposite for photocatalytic degradation of congo red dye", *Cleaner Chemical Engineering*, vol. 1, p. 100003, 2022.
doi: 10.1016/j.clce.2022.100003.
- [11] Y. Yang, K. Liu, F. Sun, Y. Liu, and J. Chen, "Enhanced performance of photocatalytic treatment of Congo red wastewater by CNTs-Ag-modified TiO₂ under visible

- light", *Environmental Science and Pollution Research*, vol. 29, pp. 15516-15525, 2022.
doi: 10.1007/s11356-021-16734-w.
- [12] P. L. Hariani, M. Said, S. Salni, N. Aprianti, and Y. A. L. R. Naibano, "High efficient photocatalytic degradation of methyl orange dye in an aqueous solution by $\text{CoFe}_2\text{O}_4\text{-SiO}_2\text{-TiO}_2$ magnetic catalyst", *Journal of Ecological Engineering*, vol. 23 (1), pp. 118-128, 2022.
doi: 10.12911/22998993/143908.
- [13] Z. Benredjem, K. Barbari, I. Chaabna, S. Saaïdia, A. Djemel, R. Delimi, S. Douas, and K. Bakhouch, "Comparative investigation on the removal of methyl orange from aqueous solution using three different advanced oxidation processes", *International Journal of Chemical Reactor Engineering*, vol. 19 (6), pp. 597-604, 2021.
doi: 10.1515/ijcre-2020-0243.
- [14] A. L. Butt, J. K. Mpinga, and S. M. Tichapondwa, "Photo-Fenton oxidation of methyl orange dye using South African ilmenite sands as a catalyst", *Catalysts*, vol. 11, p. 1452, 2021.
doi: 10.3390/catal11121452.
- [15] N. E. Chadi, S. Merouani, O. Hamdaoui, M. Bouhelassa, and M. Ashokkumar, " H_2O_2 /periodate (IO_4^-): A novel advanced oxidation technology for the degradation of refractory organic pollutants", *Environmental Science: Water Research & Technology*, vol. 5 (6), pp. 1113-1123, 2019.
doi: 10.1039/C9EW00147F.
- [16] N. E. Chadi, S. Merouani, O. Hamdaoui, M. Bouhelassa, and M. Ashokkumar, "Influence of mineral water constituents, organic matter and water matrices on the performance of $\text{H}_2\text{O}_2/\text{IO}_4^-$ -advanced oxidation process", *Environmental Science: Water Research & Technology*, vol. 5 (11), pp. 1985-1992, 2019.
doi: 10.1039/C9EW00329K.
- [17] Y. Sukhatskiy, M. Sozanskyi, M. Shepida, Z. Znak, and P. R. Gogate, "Decolorization of an aqueous solution of methylene blue using a combination of ultrasound and peroxate process", *Separation and Purification Technology*, vol. 288, p. 120651, 2022.
doi: 10.1016/j.seppur.2022.120651.
- [18] J. Oakes, and P. Gratton, "Kinetic investigations of the oxidation of Methyl Orange and substituted arylazonaphthol dyes by peracids in aqueous solution", *Journal of the Chemical Society, Perkin Transactions*, no. 2, pp. 2563-2568, 1998.
doi: 10.1039/A807272H.
- [19] Z. O. Znak, Y. V. Sukhatskiy, O. I. Zin, S. V. Khomyak, R. V. Mnykh, and A. V. Lysenko, "The decomposition of the benzene in cavitation fields", *Voprosy khimii i khimicheskoi tekhnologii (Issues of Chemistry and Chemical Technology)*, vol. 1 (116), pp. 72-77, 2018.
- [20] M. Mozurkewich, and S. W. Benson, "Negative activation energies and curved Arrhenius plots. 1. Theory of reactions over potential wells", *The Journal of Physical Chemistry*, vol. 88, no. 25, pp. 6429-6435, 1984.
doi: 10.1021/j150669a073.
- [21] D. Baena-Baldiris, A. Montes-Robledo, and R. Baldiris-Avila, "*Franconibacter* sp., 1MS: A new strain in decolorization and degradation of azo dyes Ponceau S Red and Methyl Orange", *ACS Omega*, vol. 5, pp. 28146-28157, 2020.
doi: 10.1021/acsomega.0c03786.
- [22] S. Yang, R. Jin, Z. He, Y. Qiao, S. Shi, W. Kong, Y. Wang, and X. Liu, "An experimental study on the degradation of methyl orange by combining hydrodynamic cavitation and chlorine dioxide treatments", *Chemical Engineering Transactions*, vol. 59, pp. 289-294, 2017.
doi: 10.3303/CET1759049.
- [23] S. Srinivasan, and S. K. Sadasivam, "Biodegradation of textile azo dyes by textile effluent non-adapted and adapted *Aeromonas hydrophila*", *Environmental Research*, vol. 194, p. 110643, 2021.
doi: 10.1016/j.envres.2020.110643.

Yu. V. Sukhatskiy, Ph. D.,
 e-mail: yurii.v.sukhatskiy@lpnu.ua
M. V. Shepida, Ph. D.,
 e-mail: maryana_shepida@ukr.net
M. A. Sozanskyi, Ph. D.,
 e-mail: martyn.a.sozanskyi@lpnu.ua
Z. O. Znak, Dr. Tech. Sc., Professor,
 e-mail: znak_zo@ukr.net
S. V. Khomyak, Ph. D., Senior Researcher,
 e-mail: semen.v.khomyak@lpnu.ua
 Lviv Polytechnic National University
 S. Bandery st. 12, Lviv, 79013, Ukraine

KINETICS OF OXIDATIVE DEGRADATION OF METHYL ORANGE USING THE “SONOPEROXATE” PROCESS

Disadvantages of photocatalytic and Fenton-like advanced oxidation processes of dyes are high energy consumption and long duration, which is due to the formation of only one main reactive species – hydroxyl radicals. To increase the degradation rate of azo dyes, in particular methyl orange (MO), it is proposed for the first time to use the “Sonoperoxate” advanced oxidation process (ultrasound/H₂O₂/KIO₄), during the implementation of which a complex of reactive species is formed, which includes both radicals (iodyl, periodyl, hydroxyl) and a highly active compound of molecular type (singlet oxygen – ¹O₂). Rational conditions for the oxidative degradation of MO are established: the initial concentration of MO in its aqueous solution is 76.5·10⁻⁶ mol/L; the molar ratio of MO:H₂O₂:KIO₄=1:50:10; the pH value of the reaction medium – 3.00; the temperature of the medium – 293 K; the specific power of ultrasonic cavitation treatment is 68.0 W/L. The oxidative degradation rate constant of MO, which under rational conditions is equal to 2.883·10⁻³ s⁻¹, is determined. The negative value (-19472 J/mol) of the effective activation energy of the oxidative degradation of MO using the “Sonoperoxate” advanced oxidation process is associated with the multi-stage nature of this process. The synergistic effect, which is due to the combination of the action of ultrasonic cavitation and the H₂O₂/KIO₄ process, is characterized by the value of the synergistic coefficient, which is equal to 1.62. Based on the results of UV/Vis and IR spectroscopy, the destruction of the azo bond in molecules of MO and the dye degradation are confirmed. According to the results of the identification of the functional groups of intermediates by the IR spectroscopy method, a possible pathway of oxidative degradation of MO using the “Sonoperoxate” process is proposed, which sequentially includes breaking the azo bond, demethylation of primary amines, deamination of amino derivatives of benzene, opening of the benzene ring with the formation of linear compounds, and their mineralization.

Keywords: azo dyes, oxidative degradation, rate constant, ultrasonic cavitation, periodate, UV/Vis spectroscopy, IR spectroscopy.